2016

Reliable and Efficient Transmission of Signals: Coding Design, Beamforming Optimization and Multi-Point Cooperation

Yang Liu
Lehigh University

Reliable and Efficient Transmission of Signals: Coding Design, Beamforming Optimization and Multi-Point Cooperation

by

Yang Liu

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Electrical Engineering

Lehigh University

January 2016

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

_____

Date

_____

Dissertation Advisor

_____

Accepted Date

Committee Members:

_____

Prof. Tiffany Jing Li(Chair)

_____

Prof. Zhiyuan Yan

_____

Prof. Parvathinathan Venkitasubramaniam

_____

Prof. Liang Cheng

Acknowledgements

I would like to express my deepest gratitude to my advisor Prof. Tiffany Jing Li for her guidance and support during my whole Ph.D study. I am also very grateful to my committee members, Prof. Zhiyuan Yan, Prof. Parv Venkitasubramaniam, and Prof. Liang Cheng for their help and suggestions. Without their guidance and help, I could never finish this dissertation.

I would like to thank all my friends and colleagues, especially Kai Xie, Peiyu Tan, Xingkai Bao, Xuebin Wu, Yang Yang, Chuanming Wei, Nattakan Puttarak, Phisan Kaewprapha, Feng Shi, Chen Chen, Hongmei Xie, Chenrong Xiong, Jun Lin, Xuanxuan Lv, Jiangfan Zhang, with whom I have worked together, for their encourage, help and friendship.

Last but not least, I would like to thank my parents for their love, understanding and support.

Contents

Acknowledgements iii

List of Figures ix

Abstract 1

1 Introduction **4**

1.1 Background and Motivation 4

1.2 Outline of the Dissertation 8

2 Point-to-Point Linear Analog Signal Transmission **13**

2.1 Introduction 13

2.2 Signal and System Model 16

2.3 Numerical Results 26

2.4 Conclusion 27

3 Nonlinear Analog Point-to-Point Transmission—A Family of Chaotic Pure Analog Coding Schemes Based on Baker's Map Function **30**

3.1 Introduction 30

3.2 The Baker's Map Analog Coding Scheme 34

3.2.1 Maximum Likelihood(ML) Decoding 37

3.2.2 Minimum Mean Square Error(MMSE) Decoding 40

3.2.3 Mixed ML-MMSE Decoding Scheme 41
3.2.4 Performance Analysis 43
3.3 Improvement I—Mirrored Baker's Analog Code 45
3.3.1 ML Decoding 49
3.3.2 MMSE Decoding 50
3.3.3 ML-MMSE Decoding 51
3.4 Improvement II—Single-Input(1-D) Baker's Analog Code 52
3.4.1 ML Decoding Scheme 53
3.4.2 MMSE Decoding Scheme 54
3.4.3 ML-MMSE Decoding Scheme 54
3.5 Simulation Results and Discussions 55
3.6 Conclusion 65
3.7 Appendix 66

4 Nonbandwidth-Expansion Precoding for Sensor Wireless Network Transformation 70

4.1 Introduction 70
4.2 System Model 72
4.3 Precoder Design 74
4.3.1 PEP Analysis 75
4.3.2 Discussion—High and Low SNR Cases 76
4.4 Detection Scheme at the Destination 82
4.4.1 Exact and Nearly Exact Detection Methods 82
4.4.2 LMMSE Round Off 83
4.4.3 Partially Nulling and Canceling Method 84
4.5 Numerical Results 87

4.6 Conclusion 88

5 Joint Transceiver Design towards MSE Minimization for Wireless Sensor Network 91

5.1 Introduction 91

5.2 System Model 95

5.3 Two-Block Coordinate Descent (2-BCD) 99

5.3.1 $(P1)$: Optimizing $\mathbf{G}$ given $\{\mathbf{F}_i\}$ 99

5.3.2 $(P2)$: Optimizing $\{\mathbf{F}_i\}$ given $\mathbf{G}$ 100

5.3.3 Convergence of 2-BCD Algorithm 104

5.4 Multi-Block Coordinate Descent 105

5.4.1 Further Decoupling of $(P2)$ and Closed-Form Solution 105

5.4.2 Layered-BCD Algorithm 110

5.4.3 Essentially Cyclic $(L+1)$-BCD Algorithm 112

5.4.4 Acceleration by Approximation 115

5.5 Numerical Results 116

5.6 Conclusion 123

5.7 Appendix 124

5.7.1 Proof of Theorem 5.4.1 124

5.7.2 Proof of Theorem 5.4.3 129

6 Joint Transceiver Design towards MI Maximization for Wireless Sensor Network 134

6.1 Introduction 134

6.2 System Model 140

6.3 Algorithm Design 145

6.3.1 Jointly Optimizing $\{\mathbf{F}_i\}_{i=1}^{L}$ . . . 150
6.3.2 Cyclic $(L+1)$-BCA Algorithm . . . 154
6.3.3 Complexity . . . 161
6.4 Numerical Results . . . 162
6.5 Conclusion . . . 168
6.6 Appendix . . . 168
6.6.1 Proof of Theorem 6.3.2 . . . 168

7 Joint Transceiver Design towards SNR Maximization for Wireless Sensor Network 174

7.1 Introduction . . . 174
7.2 System Model . . . 178
7.3 Optimal Linear Receiver . . . 182
7.4 Jointly Optimizing Beamformers at Sensors . . . 183
7.4.1 Solving $(P1)$ by Semidefinite Relaxation . . . 185
7.4.2 Iteratively Solving $(P1)$ . . . 191
7.4.3 Convergence and Complexity . . . 195
7.5 Multiple Block Framework to Maximize SNR . . . 197
7.5.1 One-Shot SDR-Rank-Reduction Method . . . 198
7.5.2 Iterative Method . . . 199
7.6 Numerical Results . . . 203
7.7 Conclusion . . . 208
7.8 Appendix . . . 210
7.8.1 Proof of Lemma 7.4.1 . . . 210
7.8.2 Proof of Lemma 7.4.2 . . . 211
7.8.3 Proof of Lemma 7.4.3 . . . 213

7.8.4 Proof of Lemma 7.4.4 214
7.8.5 Proof of Theorem 7.4.2 215
7.8.6 Proof of Theorem 7.5.1 218
7.8.7 Proof of Theorem 7.5.2 219

8 Conclusion **223**

Bibliography **225**

References **225**

Vita **237**

List of Figures

2.1 Group I: Random and structured unitary codes (with orthonormal rows in $\boldsymbol{G}$). 28

2.2 Group II: random codes with orthogonal but non-equal-energy rows in $\boldsymbol{G}$ 29

2.3 Group III: random codes (with independent rows in $\boldsymbol{G}$) 29

3.1 Partition and Itinerary— Left of (a): When $N = 1$, itinerary $\mathbf{s}$ has just one bit, $+1$ or -1; Right of (a): For general $N = n{+}1$, itinerary $\mathbf{s}$ has 2^n patterns. Each specific pattern $\mathbf{s}_j$ corresponds to one segment (cell) $C_{\mathbf{s}_j}$ of the feasible region; (b): The feasible region $[-1, +1]^2$ is partitioned into 2^{N-1} cells, with each cell $C_{\mathbf{s}_j}$ corresponding to one specific itinerary pattern $\mathbf{s}_j$. The parameters in the affine representation of the codewords and the endpoints of the cell can be determined once the itinerary $\mathbf{s}_j$ is given. 38

3.2 MSE performance of different decoding algorithms for baker's dynamic system 44

3.3 MSE performance for x_0 and y_0 of baker's system 46

3.4 1-D baker's dynamic encoding system 53

3.5 MSE of different decoding algorithms for mirrored baker's analog code 56

3.6 MSE performance of different decoding algorithms for single-input baker's analog code 56
3.7 MSE performance of tent map code, mirrored baker's map code and single-input baker's map code(coderate 1/10) 57
3.8 BER of itinerary bits for tent map code, mirrored baker's map code and single-input baker's map code($N = 5$) 60
3.9 Approximated OPTA, SDR of Mirrored Baker's Map Code and Shannon-Kotel'nikov Spirals with Different Parameters 61
3.10 Analog signals transmission—analog system vs digital system 64

4.1 The Sensor-Center Wireless Communication System 72
4.2 SNR at Sensor: 8dB, 4×4 unitary matrix 89
4.3 SNR at Sensor: 20dB, 4×4 ML and $4/8$ Partially Canceling and Nulling 89
4.4 SNR at Sensor: 20dB, 5×5 ML and $5/10$ Partially Canceling and Nulling 90

5.1 Multi-Sensor System Model 96
5.2 MSE Performance of 2-BCD v.s. Layered $(L+1)$-BCD (with 2 inner-loop iterations) Algorithms 118
5.3 MSE Performance of 2-BCD v.s.$(L+1)$B-FG Algorithms 118
5.4 MSE Performance of 2-BCD v.s. Approximate $(L+1)$B-FG 119
5.5 MSE Performance of 2-BCD v.s. Proximal $(L+1)$B-FG 119
5.6 MSE Performance of 2-BCD v.s. Approximate-Proximal $(L+1)$B-FG . 120
5.7 MSE Itineraries of 2-BCD v.s. $(L+1)$B-FG Algorithm 121
5.8 MSE Itineraries of 2-BCD v.s. Proximal $(L+1)$B-FG Algorithm 121
5.9 MSE Itineraries of 2-BCD v.s. Approximate-Proximal $(L+1)$B-FG Algorithm 122

6.1 Multi-Sensor System Model 141
6.2 Heterogenous Test Case: 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Numbers of Iterations. 164
6.3 Homogeneous Test Case:3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Numbers of Iterations 165
6.4 Scalar Source Signal Case: 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Numbers of Iterations. 165
6.5 Heterogenous Test Case: Optimizing Generalized MI by 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Initial Points 166
6.6 Homogeneous Test Case: Optimizing Generalized MI by 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Initial Points 167

7.1 Models for Oversampling or Cluster-Based WSN 178
7.2 Average SNR Obtained by SDR Based 2BCA Algorithm and SOCP Based 2BCA Algorithm 204
7.3 Average SNR Obtained by SDP Based 2BCA Algorithm and Multiple BCA Algorithm 205
7.4 Convergence with Different Initials: SDR Based 2BCA Algorithm v.s. SOCP Based 2BCA Algorithm 206
7.5 Convergence with Different Initials: SDR Based 2BCA Algorithm v.s. Multiple BCA Algorithm 206
7.6 Complexity of Algorithms with Respect to K 207
7.7 Complexity of Algorithms with Respect to L 207

Abstract

This dissertation focuses on reliable and efficient signal transmission strategies in classical two-point and also multi-point communication systems. Although the digital signal processing and communication technologies have been developed very successfully nowadays, the optimal non-digitized communication schemes has always been an attracting and open issue for the past decades. At the same time, enormous amount of communication networks have emerged in a wide range of fields. Its wide applications have also arisen numerous signal transmission design problems to satisfy different requirement and constraints for various system. This dissertation is devoted to some facets of these problems. Specifically, the contribution of this dissertation is as follows.

The first contribution is that the linear analog coding schemes' performance limit and optimal codes have been obtained. Under the general model of additive white Gaussian noise(AWGN) channel and mean square error(MSE) performance metric, the optimal linear analog codes under maximum likelihood(ML) and linear minimum mean square error(LMMSE) criteria are studied. The performance limits using these two decoding schemes have been obtained and they lead to identical optimal linear analog codes—*unitary codes*.

The second contribution is that we propose a novel nonlinear analog coding schemes based on chaotic dynamic systems—baker's dynamic system. Under the general AWGN channel model, various decoding algorithms have been researched, including the minimum mean square error(MMSE) decoding algorithm, maximum likelihood(ML) decoding algorithm and ML-LMMSE algorithms. MMSE algorithm provides optimal decoding performance in MSE, but it also requires prior knowledge and highly nonlinear computation operations. ML and ML-MMSE algorithms are sub-optimal decoding

schemes, which do not require knowledge of source's distribution and only involve linear computation. Based on the careful examination of the baker's dynamic system's performance limit, two improving schemes, mirrored baker's dynamic systems and one input baker's system, are proposed. The improvement schemes effectively depress the threshold effect of the original system and outperform the other existing chaotic analog codes systems in literature.

A third contribution is that we consider the precoding design for single sensor with single antenna by exploiting signal space diversity. By analyzing pairwise error probability, we discuss precoder design criterion. Besides, suboptimal decoding algorithms with low complexity are researched. A kind of partially nulling and canceling(PNC) algorithm is proposed. Extensive numerical results show the proposed PNC algorithm can achieve better bit error rate(BER) performance with even lower computation complexity.

Last but not least, the final contribution is the research on joint transceiver design in centralized wireless sensor networks. A wide range of commonly used performance measures, including MSE, mutual information(MI) and signal to noise ratio(SNR), have been taken into consideration. Under the setup of complex wireless sensor networks involving numerous variables and constraints, the joint transceiver design problems generally have highly non-convex optimization objective and extremely hard. Instead of solving these hard problems in one shot, we adopt the methodology of block coordinate descent(BCD) methods, to solve these problems in an iterative manner.By possibility necessary equivalent transformation of the original problems, we partition the whole variable space into multiple groups and each time the objective is optimized with respect to only one group of variables with the others being fixed. For the MSE, MI and

SNR optimization problems, we decompose each them into multiple convex subproblems and by analyzing the optimality conditions, most of these subproblems' closed form solutions are obtained, which significantly decrease the complexity of proposed algorithms. Besides that, convergence characteristics are also of great concerns and carefully examined. Numerical results fully verify our proposed algorithms.

Chapter 1

Introduction

1.1 Background and Motivation

Nowadays we are living in a digital communication technology era. After half century's rapid evolution, the digital communication technologies, including the digital error correction code(DECC) technologies, have been developed so successfully that their application have reached almost every facet of our modern life. The Shannon's fundamental source-channel separation theorem tells us that we can transmit signals without losing optimality by separately source-coding and channel-coding the transmitted signal. This has also cornerstoned the typical transmitter structure in the modern digital communication system—the signals are quantized, compressed, coded(perhaps modulated) and then transmitted. The digital channel coding technology has been extensively researched during the last two decades and some extremely powerful Shannon limit approaching codes, like Turbo code and low-density parity-check(LDPC) code have been

found and utilized in practice, which significantly improves the reliability of communication system.

Although the digital communication and error correction code technologies have achieved great success, like very coin having two faces, they have intrinsic drawbacks. First, many source signals in real world are born analog, like sound, light, pressure, temperature and so on. To suit these analog signals to digital communication systems, quantization must be performed. This will inevitably introduce permanent information loss. Although utilizing sufficient number of digits to represent a real valued signal can suppress the quantization noise, it extensively extends the system bandwidth, let alone the further bandwidth extension required by the subsequent channel encoding. Second, although the powerful digital error correction codes can approach Shannon's limit, they exhibit a kind of *threshold effect*. In fact all the digital error correction codes are signal-to-noise-ratio(SNR) orientated. This means that when the SNR at the receiver is lower then some threshold, the performance is usually very poor. However, once the SNR is above the threshold, the performance improves drastically within a narrow SNR range and will hardly meliorate for additional transmission power. This is non-energy-efficient and the system performance does not degrade gracefully.

Compared to the fully developed digital communication schemes, we are interested in its counterpart—analog coding and communication system. Analog coded system transmits continuous signals, which has no quantization noise and performance degrades gracefully. Actually efficient analog transmission scheme has always been an open and active issue. Marshall and Wolf propose the term *analog codes* in their paper [1, 2] to protect signals in real or complex domains. The work [15] shows that for bandwidth non-expansion communication, linear analog communication scheme is optimal. The

recent advance in analog communication systems can be found in the reference [19, 20, 22, 25, 26, 76].

For the study of analog coding system, many kinds of coding schemes exist. Generally speaking, these schemes can be divided into two kinds—linear and non-linear schemes. From engineering perspective, linear systems are always preferred due to their easiness in analysis and implementation. So part of this dissertation is devoted to the performance limit of linear analog codes and their optimal coding schemes.

For the nonlinear analog coding schemes, the optimal encoding method is still unknown. Some specially structured encoding schemes are proposed and analyzed in [21, 22, 25, 76]. One interesting potential scheme is to construct analog codes through chaotic dynamic systems. Chaotic dynamic system is a special kind of dynamic system whose state transfer is governed by functions with fast divergence feature. In a short word, for chaotic dynamic systems, any tiny perturbation of the input will result in significantly different output in a short time. This characteristic is also popularly known as the famous *butterful effect*. The seminal paper [28] first notice chaotic dynamic systems' interesting feature and applied it to constructing analog codes. In this dissertation, a new kind of nonlinear analog code based on chaotic dynamic system has been proposed and analyzed, including its performance limit, optimal and suboptimal decoding schemes, which outperforms the performance of the original chaotic analog code in [28].

Besides the pursuit of the fundamental problem of optimal analog coding scheme in point-to-point communication system, this dissertation also dedicates some parts to different kinds of optimal analog coding schemes in multi-point communication network. In recent years, due to explosive demands of information sharing, wireless communication is required in more and more scenarios where multiple points participate in signal

transmission simultaneously and cooperatively. Thus communication networks emerge almost everywhere. To name a few, cloud computing, cloud storage and internet of things have already changed our life in many ways. Among them, wireless sensor network(WSN) is a typical and attracting example. Wireless sensor networks have found their applications in a extremely wide ranges including environment monitoring, battle field surveillance, manufacture control and so on [48, 49, 51].

Wireless sensor network is comprised of multiple spatially distributed sensors. Each sensor harvests information from its neighborhood environment and wirelessly communicates with other peers. Wireless sensor network has extremely high flexibility to accommodate to different tasks, including its network organization strategy, communication protocol, routing strategy and signal processing algorithm. On the other side, this also raises great challenges the to network design. Above all, the communication reliability and efficiency in wireless sensor network is in the first place. One standard method is the transceiver(also known as beamformer or precoder) technology, which evoked great attention and extensive research during the past decade. Essentially, the transceiver design problem is linear analog coding problem. Applying transceivers in wireless networks and taking into account their characteristics like cooperation, low latency and limit size(or number of antenna), various meaningful and interesting transceiver design problems arise. This dissertation considers several facets of the transceiver design in wireless sensor network.

In many applications, due to the low cost or small size requirement, it is impossible to equip sensors with multi-antennas. Although error correction codes, like turbo or LDPC codes, can improve performance, its high latency makes it inhibitable applied to centralized wireless network. What is more, in the network with numerous sensors,

bandwidth is crucially concerned. Thus an non-bandwidth-extension precoding scheme for signal antenna sensor is desirable. Inspired by the signal space diversity proposed in [45], we devote one chapter to discuss its precoding design criterion and efficient decoding methods in practice.

Besides the single sensor precoding scheme mentioned above, we also focus on joint transceiver design in centralized wireless sensor network. In a centralized wireless sensor network, there exist nodes called fusion centers(FC), which collect data from its neighboring sensors and perform further processing and fusion. We need to focus on various performance measures due to the different natures of tasks. Standard performance metrics include mean square error(MSE), mutual information(MI), signal-to-noise ratio(SNR), which describe the effectiveness of the communication from different perspectives. The joint transceiver design problems towards different criteria, are usually very difficult problems due to its large number of variables and constraints and efficient solutions are highly desirable. Several chapters of this dissertation will focus on these topics.

1.2 Outline of the Dissertation

This dissertation is organized as follows:

Chapter 2 focuses on optimal linear analog coding schemes. Under the model of additive white Gaussian noise(AWGN) channel and the performance metric of mean square error(MSE), two optimal decoding schemes have been considered—maximum likelihood(ML) and linear minimum mean square error(LMMSE) methods. The perfor-

mance limits under these two criteria have been established and the optimal encoding schemes which can achieve these performance bounds have been identified. It is proved in this chapter that under ML and LMMSE criteria, their separate optimal MSE bounds can be simultaneously achieved by linear analog codes having (parts of) unitary encoding matrices—which we named as *unitary codes*. Extensive numerical results verify what we have found.

Chapter 3 concentrates on the problem of constructing nonlinear analog coding schemes based on chaotic dynamic systems. A novel nonlinear analog encoding scheme, baker's dynamic system is proposed, which is constructed from the baker's map, a two-dimension chaotic function. Under the general AWGN channel model, various decoding algorithms are throughly studied. The minimum mean square error(MMSE) decoding algorithm has been derived, which provides optimal performance by means of MSE. Noting that MMSE decoding algorithm requires prior knowledge of probability density function(pdf) of the source signal and involves highly nonlinear computations which are quite computation demanding, we proceed to develop two suboptimal algorithms—maximum likelihood(ML) decoding algorithm and ML-LMMSE algorithms. These two algorithms do not need prior knowledge of the source signal and only need linear operations during the whole decoding procedure. Numerical results suggest that baker's dynamic system has unsatisfying performance. Carefully examining the baker's dynamic system's performance limit via Cramer-Rao bound(CRB) reveals that bottleneck of performance lies in the unbalanced protection from its two branches. Bases on this insight, two improving schemes are proposed—mirrored baker's dynamic systems and one input baker's system. These two improvements effectively depress the threshold effect of the original system and significantly outperform the prototype chaotic analog encoding system proposed in [28].

The remaining chapters 4-7 of this dissertation are dedicated to problems related to wireless sensor networks. In chapter 4, we consider the precoding design for single sensor with single antenna by exploiting signal space diversity. By analyzing pairwise error probability(PEP), we discuss precoder design criteria. Additionally, taking into consideration the low latency requirement for the fusion center in the wireless sensor networks, suboptimal decoding algorithm with low complexity is considered. A kind of partially nulling and canceling(PNC) algorithm is proposed. Extensive numerical results show that the proposed algorithm presents a good compromise between the decoding complexity and bit error rate(BER) performance.

Chapter 5 focuses on transceiver design problem minimizing MSE in a centralized wireless sensor network. Based on the fact that the original problem is highly non-convex and difficult, we adopt the methodology of block coordinate descent(BCD) method, whose main philosophy is to partition the whole variable space into multiple groups and each time the objective is optimized with respect to only one group of variables with the others being fixed. Thus the original difficult problem can be addressed by iteratively solving a sequence of easy subproblems. In this chapter we first propose a 2-BCD method. We show that one subproblem is a minimum mean square error(MMSE) problem with a closed form solution given by Wiener filter. The other subproblem is proved to be convex with respective to all sensor beamformers jointly and can be reformulated as a second order cone programming(SOCP) problem, which can be efficiently solved by standard convex solver. Based on that, we further decompose the second subproblem into multiple atom problems with each atom problem dealing with the transmitter of only one individual sensor, whose closed form solution is obtained and consequently the complexity is decreased. The convergence of these BCD based algorithms are carefully examined and extensive numerical results are provided to verify

their performance.

In chapter 6, we consider the problem of maximizing mutual information in the centralized wireless sensor network discussed in chapter 5 under Gaussian signaling assumption. The objective is still highly non-convex. To solve this problem, the BCD methodology is hard to directly apply to. Inspired by the seminal idea of weighted minimum mean square error (WMMSE) method in [60, 73], we introduce two complicating intermediate variables—weight matrix and a virtual FC receiver as intermediate variables to make the problem more friendly to BCD algorithm. Based on that, we first decompose the MI problem into three subproblems—one subproblem to update the virtual FC receiver, one subproblem to update the weight matrix and the third subproblem to jointly optimize the entire beamformers of all sensors. The Karush-Kuhn-Tucker(KKT) conditions have been examined and we manage to prove that the solutions of this 3-BCD algorithm are KKT points. Based on that, by noticing that the third subproblem is actually similar to the one appeared in chapter 6 for MSE optimizing problem, we further decompose this subproblem into multiple smaller problems and closed form solution to each of them is available.

In chapter 7 we consider the problem of maximizing signal to noise ratio(SNR) in the same wireless sensor network model established in chapter 5 and 6. This problem has a quadratic fractional objective function and thus difficult. Still utilizing the block coordinate descent method, we decompose the original problem into subproblems optimizing individual transmitter or FC receiver alone. For the receiver optimization, the subproblem can be easily solved by generalized eigenvalue decomposition. For the transmitter optimization subproblem, we further transform the quadratic fractional problem into a sequence of quadratic problems which are still nonconvex. By use of S-lemma [68]

and recent results in rank-one matrix decomposition [86], each transformed nonconvex quadratic problem can be addressed by solving a semidefinite programming(SDP) problem followed by rank-one decomposition. Numerical results are provided to verify our proposed algorithms.

Chapter 8 concludes this dissertation.

Chapter 2

Point-to-Point Linear Analog Signal Transmission

2.1 Introduction

Linear digital error correction coding technology has been developed so successfully that its application has entered almost every corner of today's communication, computing and storage systems. This chapter studies its analog counterpart: linear analog codes, or, transformation through analog matrices. Incepted independently by Marshall and Wolf in the eighties, analog codes are also termed *real number codes* [1] [2], as the support domain of these codes are the real or the complex fields, rather than the discrete finite fields as in the case of digital codes. An (N, K) linear analog is defined by its generator matrix $\boldsymbol{G}_{K\times N} \in \mathcal{R}^{K\times N}$, and encodes a length-$K$ real or complex vector to a length-N real or complex vector via matrix multiplication.

One major motivation for considering coding in the analog domain stems from the fact that many real-world signals are by nature analog. Examples include sound, color, and various geo-, bio-, and medical-signals captured by the sensing systems.

The renowned sampling theory states that it is possible to turn a signal that is continuous in time to one that is discrete in time (i.e. Nyquist sampling) without any information loss. In comparison, quantization, an essential process in analog-to-digital (A/D) conversion which makes a signal that is continuous in amplitude to one that is discrete in amplitude, inevitably introduce permanent information loss due to rounding. If inadequate levels (bins) are used in quantization, the system performance will be dominantly deteriorated by this granularity noise, even through subsequent digital signal processing and digital coding are performed perfectly.

On the other hand, to suppress quantization noise would in general require an increase of the quantization level, which in turn results in a significant increase in data volume.

When a real value is quantized and represented as a string of binary bits, the most significant bit and the least significant bit certainly carries very different levels of importance (consider a bank account whose actual balance of $100,001 is mistaken to $100,000 versus to $000,001!). Since random attenuation and noise corruption usually occur equal-probably to every transmitted bit in a practical communication channel, sophisticated design issues arise as how to evaluate the importance of each bit and how to balance the protection on the most/more and least/less important bits. Clearly, transmitting analog signals directly in their real-valued form provides a solution that naturally eliminates all of the above problems – provided that the analog signals can be conveyed with sufficient accuracy.

A second motivation for studying linear transform with analog matrices comes from its close relation to several of the modern-day wireless technologies. For example, discrete Fourier transform (DFT) matrices, an important class of analog matrices, have found intriguing use in space-time coding and modulation diversity (e.g. [5, 6]). A subset of DFT matrices also constitute the analog version of BCH codes and Reed-Solomon (RS) codes, and the later achieves the same singleton bound performance as their celebrated digital counterpart [3] [4] [7]. In particular, it has been shown that analog RS codes can be exploited to effectively combat the peak-to-average-power ratio (PAPR) issue in orthogonal frequency division multiplexing (OFDM) systems [8] [9].

The focus of this chapter is to analyze analog codes, establish their performance limits, and identify best practices. Existing research on analog codes has primarily focused on special classes of analog matrices (such as discrete cosine/sine transform (DST/DST) matrices, and discrete Fourier transform matrices), and on special decoding/detection algorithms (such as the Berlekamp-Massy algorithm and the Forney algorithm). Further, almost all the performance evaluation is on a special type of communication channel known as the *pulse channel*. A pulse channel is the analog counterpart of an erasure channel, namely, an arbitrary transmitted (analog) signal will either encounter an additive pulse noise (of arbitrary amplitude) or gets across the channel perfectly intact.

Since pulse channels are not a common channel model, this chapter considers the more realistic model of additive white Gaussian noise (AWGN). We study general analog matrices/codes, and general decoding methods. Hamming errors, which denote the number of elements that differ, is typically used to evaluate the performance of a digital system. When signals take continuous real values, two vectors may differ in every element, but differ very minorly, or, they may differ only in one element, but the differ-

ence is huge. Hence, instead of Hamming errors, we follow the convention and use the mean square error (MSE) as the distortion metric. We evaluate two classes of optimal detectors, the maximum-likelihood (ML) detector and the linear minimum mean square error (LMMSE) detector, and establishes the respective performance lower bound. For the same code, we show that LMMSE detector performs better than ML detector, but the gain is most noticeable at low to medium signal-to-noise ratio (SNR). We further characterize the optimal codes that achieve the best performance under each detector, and show that the codes designed for LMMSE detector also achieves optimality when ML detector is used, but the inverse is not true. Finally, we identify the *unitary codes*, a special class of analog codes which subsume DCT/DST and DFT codes as their instances, as the best analog codes that simultaneously achieve both lower bounds with equality. The analytical results are verified by extensive simulations.

2.2 Signal and System Model

This section discusses the signal model and encoding procedure of linear analog coding system, on which the following sections are based. We use bold fonts to denote vectors and matrices, and use regular fonts to denote scalars. Vectors are by default column vectors. The superscript ^{T} denotes the normal vector/matrix transpose, and the superscript ^{H} denotes Hermitian transpose.

The original information sequence entering the analog system is assumed to be discrete in time and continuous in value. Let $\boldsymbol{u} = (u_1, u_2, ..., u_K)^T \in \mathbb{C}^{K\times 1}$ denote the information sequence. To make the problem clear and easy to analyze, the input signal $\boldsymbol{u}$ is assumed to satisfy the following conditions

A1) Each coordinate u_i of the input sequence $\mathbf{u}$ follows an i.i.d. distribution with probability density function $p(u)$.

A2) Any coordinate u_i of the original signal sequence $\mathbf{u}$ has zero expectation, i.e.

$$E[u_i] = \int u_i p(u_i) \mathrm{d}u_i = 0 \tag{2.1}$$

Let $D(u_i) = E[u_i^2] - E^2[u_i] = D_u - 0 = D_u$ be the average energy of each coordinate of $\mathbf{u}$.

The first condition states that the input is drawn from an i.i.d. random process, and the second states that the signal space is centered around the origin for the sake of energy efficiency.

These two conditions are not special constraints, but are commonplace in communication systems.

The linear analog codes perform a linear mapping to the input sequence $\boldsymbol{u}$. Following the convention from digital error correction code, let $\boldsymbol{G} = \{g_{ji}\} \in \mathbb{C}^{K \times N}$ be the generator matrix, where $K \leq N$. The codeword, $\boldsymbol{v} \in \mathbb{C}^{N \times 1}$, is computed through the following linear encoding procedure:

$$\boldsymbol{v} = \boldsymbol{G}^H \boldsymbol{u} \tag{2.2}$$

For the code to be meaningful, $\boldsymbol{G}$ must have K mutually independent rows (full rank). When the codeword passes through an AWGN channel, the received signal $\boldsymbol{r}$ becomes:

$$\boldsymbol{r} = \boldsymbol{v} + \boldsymbol{n} = \boldsymbol{G}^H \boldsymbol{u} + \boldsymbol{n}, \tag{2.3}$$

where the noise vector $\boldsymbol{n} \in \mathbb{C}^{N\times 1}$ follows an i.i.d. complex Gaussian distribution, whose covariance $\boldsymbol{R}_n$ is a diagonal matrix ($\boldsymbol{I}$ denotes an identity matrix):

$$\boldsymbol{R}_n = \sigma^2 \boldsymbol{I}_{N\times N}. \tag{2.4}$$

From conditions *A1)* and *A2)*, we get

$$\begin{aligned} E[\boldsymbol{v}^H\boldsymbol{v}] &= E[\boldsymbol{u}^H\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{u}] \\ &= E\left[\sum_{i=1}^{N}\Big(\sum_{j=1}^{K} g_{ji}^* u_j\Big)\Big(\sum_{j=1}^{K} g_{ji} u_j^*\Big)\right] \\ &= E\left[\sum_{i=1}^{N}\Big(\sum_{j=1}^{K} |g_{ji}|^2|u_j|^2 + \sum_{j\neq k}^{K} g_{ji}^* g_{ki} u_j u_k^*\Big)\right] \\ &= \sum_{i=1}^{N}\left(\sum_{j=1}^{K} |g_{ji}|^2 E[|u_j|^2]\right) + 0 \\ &= \sum_{i=1}^{N}\sum_{j=1}^{K} |g_{ji}|^2 D_u = \|\boldsymbol{G}\|^2 D_u, \end{aligned} \tag{2.5}$$

where the $\|\cdot\|$ denotes the Frobenius norm (2-norm) of a matrix or vector.

We consider general codes, where the choice of the generator matrix $\boldsymbol{G}$ can be arbitrary, except for a power constraint. Let E_b be the power per signal coordinate. The total transmitted energy for the codeword $\boldsymbol{v}$ satisfies:

$$E[\boldsymbol{v}^H\boldsymbol{v}] = \|\boldsymbol{G}\|^2 D_u = KE_b \triangleq C. \tag{2.6}$$

The above computation can be summarized by the following system conditions:

A3) The average transmission energy of a codeword $\boldsymbol{v}$ is a constant C, as stated in

(2.6).

A4) The noise coordinates n_i follow i.i.d. circularly symmetric complex Gaussian distribution, $n_i \sim \mathcal{CN}(0, \sigma^2)$, where $\sigma^2/2 = \sigma_x^2 = \sigma_y^2$.

Before proceeding to the discussion of the general detection/decoding technique, we first establish the mean square error as the distortion metric. Let $\hat{\boldsymbol{u}}$ denote the estimate of the original signal vector $\boldsymbol{u}$, the MSE distortion Δ is defined as

$$\Delta \triangleq \frac{E\left[\,\|\boldsymbol{u}-\hat{\boldsymbol{u}}\|^2\right]}{K}, \quad \boldsymbol{u} \in \mathbb{C}^{K\times 1} \tag{2.7}$$

Two classes of optimal detectors are considered: maximum likelihood and minimum mean square error. In the context of coding, an optimal decoder usually refers to one that performs the ML detection (i.e. most probable). However, since our signals are real-valued, and since MSE serves as the figure of merit, LMMSE becomes highly relevant.

$$\text{ML criterion}: \ \arg\max_{\hat{\boldsymbol{u}}} P(\boldsymbol{r}|\hat{\boldsymbol{u}}), \tag{2.8}$$

$$\text{MMSE criterion}: \ \arg\max_{\hat{\boldsymbol{u}}} E[\,\|\hat{\boldsymbol{u}}-\boldsymbol{u}\|^2]. \tag{2.9}$$

Lemma 2.2.1. The LMMSE detector of a linear analog code outputs the follow vector:

$$\hat{\boldsymbol{u}}_{LMMSE} = (\boldsymbol{G}\boldsymbol{G}^H + \frac{\sigma^2}{D_u}\boldsymbol{I})^{-1}\boldsymbol{G}, \tag{2.10}$$

and the resultant MSE distortion per signal coordinate is

$$\Delta_{LMMSE} = \frac{D_u}{K} tr(\boldsymbol{I} + \frac{D_u}{\sigma^2}\boldsymbol{G}\boldsymbol{G}^H)^{-1}. \tag{2.11}$$

Proof. The LMMSE detector performs a linear operation to accomplish the decoding task by minimizing the MSE criterion in (2.9). Let $\boldsymbol{A}$ be the MMSE linear decoder, which leads to the estimate $\hat{\boldsymbol{u}} = \boldsymbol{A}\boldsymbol{r}$. Substituting (2.3), (2.4) and the condition *A2)* into (2.7), the MSE distortion per signal coordinate can be calculated as:

$$\Delta = \frac{1}{K} E[(\hat{\boldsymbol{u}} - \boldsymbol{u})^H (\hat{\boldsymbol{u}} - \boldsymbol{u})], \tag{2.12}$$

$$\begin{aligned}
&= \frac{1}{K} tr\big(E[(\hat{\boldsymbol{u}} - \boldsymbol{u})(\hat{\boldsymbol{u}} - \boldsymbol{u})^H]\big), \\
&= \frac{1}{K} tr\big(E[\big(\boldsymbol{A}(\boldsymbol{G}^H \boldsymbol{u} + \boldsymbol{n}) - \boldsymbol{u}\big)\big(\boldsymbol{A}(\boldsymbol{G}^H \boldsymbol{u} + \boldsymbol{n}) - \boldsymbol{u}\big)^H]\big), \\
&= \frac{D_u}{K}\Big(tr\big((\boldsymbol{A}\boldsymbol{G}^H - \boldsymbol{I})(\boldsymbol{G}\boldsymbol{A}^H - \boldsymbol{I})\big) + \frac{\sigma^2}{D_u} tr(\boldsymbol{A}\boldsymbol{A}^H)\Big).
\end{aligned} \tag{2.13}$$

For any given generator matrix $\boldsymbol{G}$, the optimal LMMSE receiver $\boldsymbol{A}_{MMSE}$ can be determined by taking the derivative in (2.13) with respect to $\boldsymbol{A}^*$ and reducing $\frac{\partial MSE}{\partial \boldsymbol{A}^*} = 0$, where $\boldsymbol{A}^*$ denotes the conjugate of matrix $\boldsymbol{A}$. We get:

$$\boldsymbol{A}_{LMMSE} = (\boldsymbol{G}\boldsymbol{G}^H + \frac{\sigma^2}{D_u}\boldsymbol{I})^{-1}\boldsymbol{G}, \tag{2.14}$$

and, hence, the LMMSE optimal estimate becomes

$$\hat{\boldsymbol{u}}_{LMMSE} = \boldsymbol{A}_{LMMSE}\boldsymbol{r} = (\boldsymbol{G}\boldsymbol{G}^H + \frac{\sigma^2}{D_u}\boldsymbol{I})^{-1}\boldsymbol{G}\boldsymbol{r}. \tag{2.15}$$

Substituting (2.14) in (2.12) leads to the following mean square error per signal coordinate:

$$\Delta_{LMMSE} = \frac{D_u}{K} tr(\boldsymbol{I} + \frac{D_u}{\sigma^2}\boldsymbol{G}\boldsymbol{G}^H)^{-1}.$$

□

Theorem 2.2.1. Consider LMMSE detector for an arbitrary (N, K) linear analog code.

(i) The resultant MSE (per signal coordinate) is lower bounded by:

$$\Delta_{LMMSE} \geq \Delta^*_{LMMSE} = \frac{D_u}{1 + \frac{E_b}{\sigma^2}}. \tag{2.16}$$

(ii) The lower bound is achieved by such linear analog codes whose generator matrix $\boldsymbol{G}_{K\times N}$ consists of K orthogonal vectors, each having the same Frobenius norm.

Proof. From the analysis in the previous section, we know that $\boldsymbol{G}$ should satisfy (2.6). At the same time, we notice the fact that

$$\|G\|^2 = tr(\boldsymbol{G}^H\boldsymbol{G}) = tr(\boldsymbol{G}\boldsymbol{G}^H). \tag{2.17}$$

Thus the constraint in (2.6) can be equivalently expressed as:

$$tr(\boldsymbol{G}\boldsymbol{G}^H) = \frac{KE_b}{D_u}. \tag{2.18}$$

We now look for the best generator matrix $\boldsymbol{G}_{K\times N}$ that will minimize the MSE distortion (per signal coordinate) under the LMMSE detector. This problem of code design can be formulated as the following optimization problem:

$$\arg\min_{\boldsymbol{G}}: \quad \Delta_{LMMSE} = \frac{D_u}{K}(\boldsymbol{I} + \frac{D_u}{\sigma^2}\boldsymbol{G}\boldsymbol{G}^H)^{-1} \tag{2.19}$$

$$s.t. \quad tr(\boldsymbol{G}\boldsymbol{G}^H) = \frac{KE_b}{D_u} \tag{2.20}$$

According to the Hadamard's inequality, for any $N \times N$ positive semidefinite matrix $\boldsymbol{B}$, we have

$$tr(\boldsymbol{B}^{-1}) \geq \sum_{i=1}^{N} \frac{1}{\boldsymbol{B}_{ii}} \tag{2.21}$$

where $\boldsymbol{B}_{ii}$ is the i-th diagonal element of matrix $\boldsymbol{B}$. The above inequality (2.21) achieves equality if and only if $\boldsymbol{B}$ is diagonal. Notice that $\boldsymbol{G}\boldsymbol{G}^H$ is positive semidefinite. We have

$$\frac{D_u}{K} tr(\boldsymbol{I} + \frac{D_u}{\sigma^2}\boldsymbol{G}\boldsymbol{G}^H)^{-1} \geq \frac{D_u}{K} \sum_{i=1}^{K} \frac{1}{1 + \frac{D_u}{\sigma^2}[\boldsymbol{G}\boldsymbol{G}^H]_{ii}} \tag{2.22}$$

Equation (2.22) holds if and only if $(\boldsymbol{I} + \frac{D_u}{\sigma^2}\boldsymbol{G}\boldsymbol{G}^H)$ is a diagonal matrix, or, equivalently, $\boldsymbol{G}\boldsymbol{G}^H$ is diagonal. We denote the diagonal matrix $\boldsymbol{Q} = \boldsymbol{G}\boldsymbol{G}^H = diag\{q_1, q_2, \cdots, q_K\}$. The problem in (2.19) converts to:

$$\arg\min_{\{q_i\}}: \quad \Delta_{MMSE} = \sum_{i=1}^{K} \frac{D_u}{1 + \frac{D_u}{\sigma^2} q_i} \tag{2.23}$$

$$s.t. \quad tr(\boldsymbol{G}\boldsymbol{G}^H) = \sum_{i=1}^{K} q_i = \frac{KE_b}{D_u} \tag{2.24}$$

$$q_i > 0, \forall i \in \{1, 2, \cdots, K\} \tag{2.25}$$

It is now easy to see that the optimality (minimum) is achieved when

$$q_1 = q_2 = \cdots = q_K = \frac{KE_b}{KD_u} = \frac{E_b}{D_u}. \tag{2.26}$$

This suggests that the best $\boldsymbol{G}$ is one whose rows are mutually orthogonal (such that $\mathrm{GG^H}$ becomes diagonal) *and* each row has the same energy (norm). With such an optimal $\boldsymbol{G}$, the objective Δ_{LMMSE} achieves the following lower bound with equality:

$$\Delta_{LMMSE} \geq \Delta^*_{LMMSE} = \sum_{i=1}^{K} \frac{D_u}{K(1 + \frac{D_u}{\sigma^2} q_i)} = \frac{D_u}{1 + \frac{E_b}{\sigma^2}}.$$

□

Definition 2.2.1. An (N, K) *unitary codes* is a linear (analog) code whose generator matrix $\boldsymbol{G}_{K\times N}$ is obtained by deleting $(N-K)$ rows of an $N \times N$ square unitary matrix. The $(N-K)$ deleted rows can be assembled to form the parity check matrix this unitary code.

Corollary 2.2.1. Unitary codes are the best linear analog code under LMMSE decoder (with respect to MSE distortion).

Proof. Since the generator matrix $\boldsymbol{G}$ of an unitary code consists of orthonormal rows, from Theorem 2, unitary codes are optimal codes under LMMSE decoder. □

Lemma 2.2.2. The ML detector of a linear analog code outputs the follow vector:

$$\hat{\boldsymbol{u}}_{ML} = (\boldsymbol{G}\boldsymbol{G}^H)^{-1}\boldsymbol{G}\boldsymbol{r}, \tag{2.27}$$

and the resultant MSE distortion per signal coordinate is given by

$$\Delta_{ML} = \frac{\sigma^2}{K} tr\big((\boldsymbol{G}\boldsymbol{G}^H)^{-1}\big) \text{ (per coordinate).} \tag{2.28}$$

Proof. Following the i.i.d. complex Gaussian distribution mentioned in condition *A4)*, the ML criterion in (2.8), which is to maximize a multivariate Gaussian distribution, can be reduced to minimizing the squared Euclidean distance (he exponential part of the Gaussian distribution becomes the cost function):

$$\begin{aligned} \arg\min_{\hat{\boldsymbol{u}}} J(\hat{\boldsymbol{u}}) &\stackrel{\Delta}{=} \|\boldsymbol{r} - \boldsymbol{G}^H\hat{\boldsymbol{u}}\|^2, \\ &= (\boldsymbol{r} - \boldsymbol{G}^H\hat{\boldsymbol{u}})^H(\boldsymbol{r} - \boldsymbol{G}^H\hat{\boldsymbol{u}}), \\ &= \hat{\boldsymbol{u}}^H\boldsymbol{G}\boldsymbol{G}^H\hat{\boldsymbol{u}} - \boldsymbol{r}^H\boldsymbol{G}^H\hat{\boldsymbol{u}} - \hat{\boldsymbol{u}}^H\boldsymbol{G}\boldsymbol{r} + \boldsymbol{r}^H\boldsymbol{r}. \end{aligned} \tag{2.29}$$

Taking the derivative, and reducing $\frac{\partial J(\hat{\boldsymbol{u}})}{\partial \hat{\boldsymbol{u}}^*} = 0$, we obtain (* denotes the complex conjugate):

$$\boldsymbol{G}\boldsymbol{G}^H\hat{\boldsymbol{u}} - \boldsymbol{G}\boldsymbol{r} = 0, \tag{2.30}$$

which leads to the ML estimate in (2.27). The second order derivative is verified to be < 0, so the result is indeed a maximum (rather than a minimum). The MSE per information signal coordinate can be computed as

$$\begin{aligned}
&\Delta_{ML} = \frac{1}{K}E[(\hat{\boldsymbol{u}} - \boldsymbol{u})^H(\hat{\boldsymbol{u}} - \boldsymbol{u})], \\
=&\frac{1}{K}tr\Big(E\big[(\hat{\boldsymbol{u}} - \boldsymbol{u})(\hat{\boldsymbol{u}} - \boldsymbol{u})^H\big]\Big), \\
=&\frac{1}{K}tr\Big(E\big[((\boldsymbol{G}\boldsymbol{G}^H)^{-1}\boldsymbol{G}\boldsymbol{r} - \boldsymbol{u})((\boldsymbol{G}\boldsymbol{G}^H)^{-1}\boldsymbol{G}\boldsymbol{r} - \boldsymbol{u})^H\big]\Big), \\
=&\frac{\sigma^2}{K}tr\Big((\boldsymbol{G}\boldsymbol{G}^H)^{-1}\Big).
\end{aligned}$$

□

Theorem 2.2.2. Consider ML detector for an arbitrary (N, K) linear analog code.

(i) The MSE distortion (per signal coordinate) is lower bounded by

$$\Delta_{ML} \geq \Delta_{ML}^* = \frac{D_u\sigma^2}{E_b}, \tag{2.31}$$

(ii) The optimal generator matrix that achieves this lower bound has orthogonal rows, and each row has the same norm (energy).

Proof. The proof here follows much the same line of proof as in Theorem 3. To minimize the MSE per signal coordinate under ML detection, we have the following code

design problem:

$$\arg\min_{\boldsymbol{G}}: \quad MSE_{ML} = \frac{\sigma^2}{K} tr\big((\boldsymbol{G}\boldsymbol{G}^H)^{-1}\big) \tag{2.32}$$

$$s.t. \quad tr(\boldsymbol{G}\boldsymbol{G}^H) = \frac{KE_b}{D_u} \tag{2.33}$$

Utilizing the equality in (2.21), we get

$$\frac{\sigma^2}{K} tr\big((\boldsymbol{G}\boldsymbol{G}^H)^{-1}\big) \geq \sigma^2 \sum_{i=1}^{K} \frac{1}{[\boldsymbol{G}\boldsymbol{G}^H]_{ii}} \geq \frac{D_u \sigma^2}{E_b}, \tag{2.34}$$

where the first equality is achieved if and only if $\boldsymbol{G}\boldsymbol{G}^H$ is diagonal, and the second equality is achieved when the rows of $\boldsymbol{G}$ also has the same norm. □

Corollary 2.2.2. For the same linear analog code over AWGN channels, the LMMSE detector always outperforms the ML detector with respect to the MSE distortion, but the gain diminishes as SNR increases.

Proof. Corollary 6 follows directly from the fact that

$$\Delta^*_{LMMSE} = \frac{D_u}{1 + \frac{E_b}{\sigma^2}} \leq \frac{D_u}{\frac{E_b}{\sigma^2}} = \Delta^*_{ML}. \quad \square$$

□

Corollary 2.2.3. Unitary codes are optimal under ML detection (as well as LMMSE detection) w.r.t. MSE distortion.

Comment: From the previous discussion, we also see a coincidence of the optimal linear codes under LMMSE detection and ML detection. (i) Although LMMSE detec-

tor always yields a better MSE performance than the ML detector, the optimal codes to both detectors are same one, namely, the generator matrices whose rows are orthonormal and have the same energy (norm). Notable examples of optimal codes are unitary codes, which subsume discrete cosine/sine transform codes and discrete Fourier transform codes as subclasses. (ii) If a generator matrix $\boldsymbol{G}$ is optimal, then its scaled version (i.e. $a\boldsymbol{G}$ is also optimal). The effective MSE distortion accounts for the scaling factor (i.e. the energy induced by the generator matrix).

2.3 Numerical Results

This section provides extensive simulations to verify and support the afore-analysis.

For comparison purpose, all the linear analog codes we use here are $(60, 30)$ codes. The input signal source is i.i.d. complex Gaussian variables following $\mathcal{CN}(0, 2)$ (unit variance for each dimension), thus we can readily calculate $D_u = 2$. Noise in channel are also circular symmetric complex Gaussian variables, whose variance is determined by specific SNR. The performance is measured by MSE per signal coordinate, presented in a log-scale $\log_2(\Delta)$. We evaluate three different code group, each consisting of five codes (generator matrices $\boldsymbol{G}$) satisfying specific properties.

1) Group I: 3 randomly-generated unitary codes ($\boldsymbol{G}$ has orthonormal rows), 1 DFT code, and 1 DCT code.

2) Group II: 5 randomly-generated $\boldsymbol{G}$, each having mutually orthogonal rows, but the rows do not have the same norm.

3) Group III: 5 randomly-generated $\boldsymbol{G}$, each having linearly independent but non-

orthogonal rows. (Note row-independence is required in order for the code to be meaningful.)

After each code is generated, a proper scaling factor is applied to ensure the average transmission energy per (information) signal coordinate is $E_b = 4$.

The performances of each group, decoded by both the ML detector and the LMMSE detector, are demonstrated in Figures 2.1 2.2, 2.3, respectively, and compared to the respective lower bounds. Our simulations confirm the analytical results that (i) the LMMSE lower bound is lower than the ML lower bound, and the two tend to converge at high SNRs; and (ii) unitary codes (Group I) are optimal and simultaneously achieve ML and LMMSE lower bounds. From Figure 2.3, we also see that (iii) an arbitrary, randomly-generated code tends to yield the same "average" performance that is somewhat far from the lower bounds. Further, it is also interesting to see that (iv) codes in Group II (Figure 2.2), whose rows are independent and orthogonal, but not having the same energy, exhibit drastically different MSE distortion with ML detection; whereas the difference becomes much less noticeable with LMMSE detection; and the gap to the theoretical bound is also smaller with LMMSE than with ML detection. This suggests that code design is more of an issue to ML detection than to MMSE detection. In general, LMMSE detection is recommended, if the noise variance is known.

2.4 Conclusion

We have analyzed the performance of linear analog codes. MSE performances lower bounds are established for ML and LMMSE detection, respectively, and unitary codes

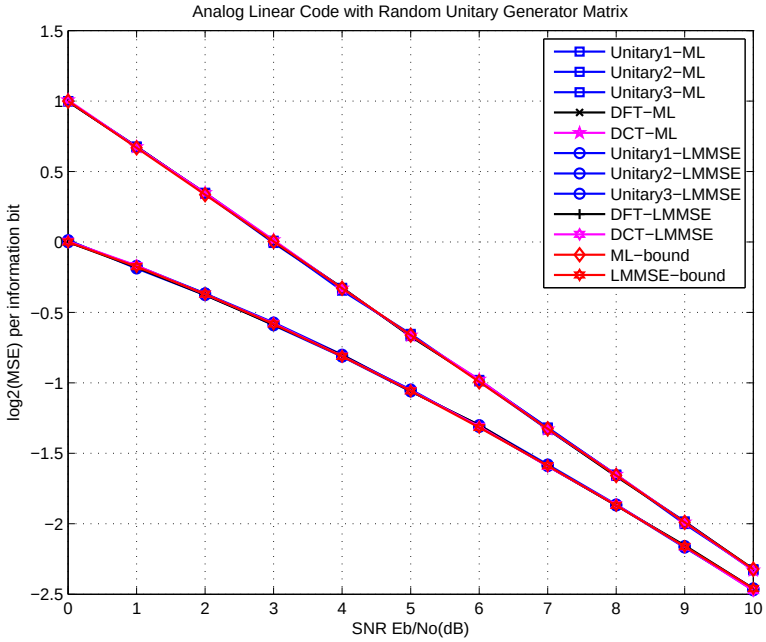

Figure 2.1: Group I: Random and structured unitary codes (with orthonormal rows in $\boldsymbol{G}$).

are identified as the optimal code that achieve both bounds. We conclude this section by emphasizing that analog codes posses unique advantages in transmitting analog signals. Further, several useful analog matrices, such as DFT matrices and DCT/DST matrices, are being actively exploited in the construction of space time codes and modulation diversity. Note that exiting systems tend to use ML detection. This study points out a possibility to reformulate the problem to one that minimizes MSE, and to use LMMSE decoder to achieve additional gains.

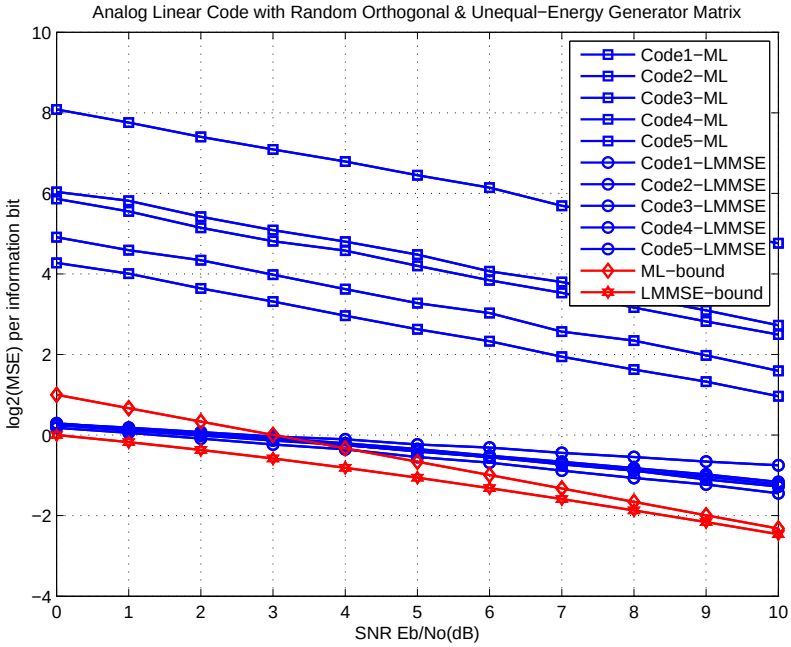

Figure 2.2: Group II: random codes with orthogonal but non-equal-energy rows in $\boldsymbol{G}$

.

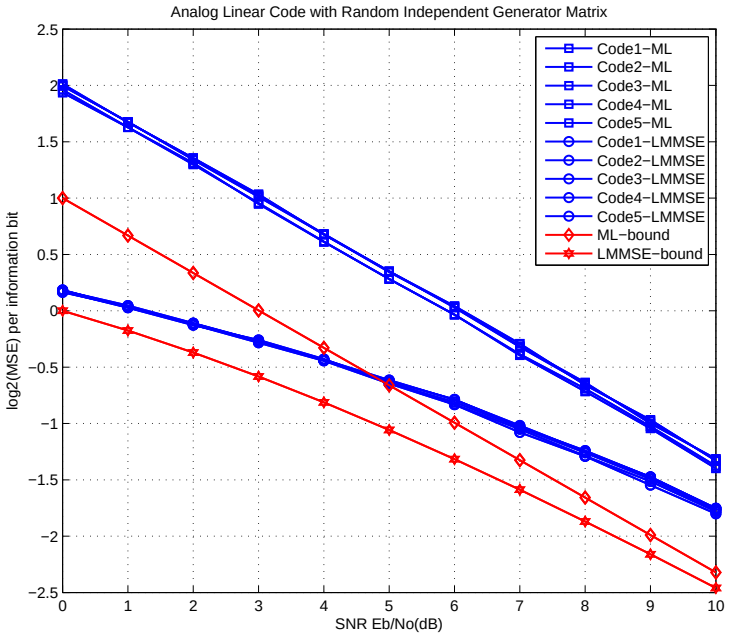

Figure 2.3: Group III: random codes (with independent rows in $\boldsymbol{G}$)

Chapter 3

Nonlinear Analog Point-to-Point Transmission—A Family of Chaotic Pure Analog Coding Schemes Based on Baker's Map Function

3.1 Introduction

Currently pervasive communication systems in practice are almost digital-based. Shannon's source-channel separation theorem has long convinced people that information can be transmitted without loss of optimality by a two-step procedure: compression and encoding. This fundamental result has laid the foundation for typical structure of modern digital communication system—the tandem structure of source coding fol-

lowed by the channel coding. Although digital communication systems have been well developed over the last decades, it has inherent drawbacks. First, to transmit continuous-alphabet sources, signals are quantized, which introduces permanent loss in information. Second, to precisely represent real-valued signals via digits, the bandwidth is usually expanded. Moreover the subsequent channel coding procedure makes transmission further bandwidth-demanding. Third, the digital error correction codes are highly signal-to-noise-ratio(SNR) dependent. Take turbo and low-density parity-check(LDPC) codes as examples. When the receiving SNR is under some threshold value, the decoding performance is usually very poor. On the contrary, once SNR exceeds this threshold, their bit error ratio(BER) falls down drastically in a narrow SNR range(waterfall region). This ungraceful degradation in performance can cause problems in applications. An typical scenario is the broadcasting system, where the SNR for different receivers can vary over a large range. At the same time the digital error correction codes are not energy efficient since more transmission power increases performance little as long as the receiving SNR is modestly above the threshold. Last but not least, digital error correction codes with satisfying performance usually require a long block-length, which introduces high latency for decoding and processing at the receiver.

In addition to the classical source-channel separate digital system, analog transmission system can serve as an alternative solution to data transmission. Analog system has advantages over its pure digital peers—it does not introduce the granularity noise and its performance evolves gracefully with SNR. Most of the analog transmission systems ever presented in literature are joint source channel coding(JSCC) systems, where compression and encoding are performed in one step and signals are in pure analog or hybrid-digital-analog(HDA) form. The study of analog communication can date back to the papers [12–15]. Reference [15] shows that direct transmission of Gaussian source

over additive white Gaussian noisy(AWGN) channel with no bandthwidth expansion or compression is optimal. For bandwidth expansion case, [16] obtains the result that the fastest decay speed of mean square error(MSE) cannot be better than square inverse of SNR. Although until now, no practical schemes have been found to achieve this decaying speed. In [17, 18], the optimal linear analog codes are treated. Design of practical nonlinear analog coding schemes has always been an open issue. Some interesting paradigms have been found. [19] and [20] discuss numerical-based analog signal encoding schemes. [21] proposes a class of analog dynamic systems constructed by first order derivative equations, which generate algebraic analog codes on torus or sphere. [22], [25] and [26] study the design of Shannon-Kotel'nikov curve. The minimum mean square error decoding schemes for Shannon-Kotel'nikov analog codqes and its modified version combined with hybrid digital signals are discussed in [76] and [24].

Among the family of analog coding schemes, one special class is constructed through chaotic dynamic systems. In dynamic systems, the signal sequence is generated by iteratively invoking some predefined mapping function. To be specific, the next signal(state) is obtained by performing a mapping to the current signal(state) and the whole signal(state) sequence is initialized by the input signal. For a chaotic dynamic system, the function governing the signal generation(state transition) is chosen as chaotic functions. Chaotic functions are characterized by their fast divergence, which is more well known as the remarkable *butterfly effect*. This property means that even a very tiny difference in initial inputs will soon result in significantly different signal sequences. From the signal space expansion viewpoint, this indicates that a pair of points in source space with small distance will have a large distance in the code space. So chaotic dynamic systems can potentially entitle signals with error resistance. The seminal work [27] proposes an analog system based on tent map dynamic system and its performance is extensively

discussed in [28]. As the analysis performed in [29, 30], the drawback of tent map code is that its performance convergence Cramer-Rao lower bound(CRLB) requires very high SNR. [31] proposes an improvement scheme by protecting the itinerary of the tent map codes with digital error correction codes. However this hybrid-digital-analog scheme still suffers from the drawbacks rooted in digital error correction codes.

In this chapter we focus on a new pure analog chaotic dynamic encoding scheme, which is constructed via a two-dimensional chaotic function—baker's map. This structure is closely related to and more complicated than the one reported in [27]. The specific contributions of this chapter include: we develop various decoding methods for the baker's coding system and analyze its MSE performance. Based on that, we proceed to propose two improved coding structures and extend various decoding methods to these new structures. These proposed improvements effectively balance the protection for all source signals and has more satisfying MSE performance compared to the tent map code. We also compare our proposed analog coding scheme with the classical source-channel separate digital coding scheme, where turbo code is applied. By using equal power and bandwidth, our proposed coding scheme outperforms the digital turbo scheme over a wide SNR range.

This chapter is organized as follows: in section.3.2, the original baker's dynamic system is discussed, including its encoding structure, decoding methods and its performance analysis. Two modified chaotic systems based on baker's system are discussed in section 3.3 and 3.4, including its encoding and decoding schemes. In section 3.5 numerical results and discussions are presented and performance are discussed. Section 3.6 concluded the chapter.

In this chapter, we assume that the source signals are mutually independent and

uniformly distributed on the interval $[-1, 1]$, which is also adopted in previous works [27] and [28]. By this assumption the MMSE decoding method has closed form solution and we can compare performance with previous works. However it should be pointed out that ML decoding method does not require this condition and is applicable to signal with arbitrary distribution. We assume that the transmission channel is AWGN and the decoding methods obtained can be easily extended to block fading channel.

3.2 The Baker's Map Analog Coding Scheme

In this section, we introduce the analog encoding scheme based on baker's map function. The baker's map function, $F : [0, 1]^2 \mapsto [0, 1]^2$, is a piecewise-linear chaotic function given as follows:

$$\begin{bmatrix} x \\ y \end{bmatrix} = F(u, v) = \begin{bmatrix} 1 - 2\mathsf{sign}(u)u \\ \frac{1}{2}\mathsf{sign}(u)(1 - v) \end{bmatrix}, \quad -1 \leq u, v \leq 1. \tag{3.1}$$

The above baker's map has an close connection with the symmetric tent map function discussed in [27] and [28], which is defined as

$$G(x) = 1 - 2|x|, \quad -1 \leq x \leq 1. \tag{3.2}$$

Although the symmetric tent map in (3.2) is non-invertible, once the $\mathsf{sign}(x)$ is given, its value can be determined. The "inverse" symmetric tent map function with the sign s of x given is $G_s^{-1}(y) = s\frac{1-y}{2}$. Comparing the baker's map and the symmetric tent

map functions, the baker's map can be alternatively defined via the symmetric tent map function as follows

$$\begin{bmatrix} x \\ y \end{bmatrix} = F(u,v) = \begin{bmatrix} G(u) \\ G^{-1}_{\mathsf{sign}(u)}(v) \end{bmatrix}, \quad -1 \le u, v \le 1. \tag{3.3}$$

Based on the baker's map function above, a dynamic analog encoding scheme can be performed. For a pair of independent source $(x_0, y_0) \in [-1,1]^2$, a chaotic signal sequence is generated by repeatedly invoking baker's mapping, i.e.

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = F(x_n, y_n), \quad n = 0, 1, \cdots, N-2, \tag{3.4}$$

where N is the bandwidth expansion. This sequence can be viewed as a rate-$1/N$ analog code with x_0 and y_0 as continuous information "bits". In the following, we use $\boldsymbol{x} = [x_0, x_1, \cdots, x_{N-1}]^T$ and $\boldsymbol{y} = [y_0, y_1, \cdots, y_{N-1}]^T$ to denote the codewords of two input signals respectively.

An important concept about the baker's dynamic encoding system is the *itinerary*, which is defined as $\boldsymbol{s} = [s_0, s_1, \cdots, s_{N-2}] \triangleq [\mathsf{sign}(x_0), \mathsf{sign}(x_1), \cdots, \mathsf{sign}(x_{N-2})]$. In fact if the itinerary of the code sequence is given, x_k's and y_k's can all be expressed as affine functions of x_0 and y_0. Specifically, x_k and y_k can be represented via (x_0, y_0) in the following form

$$\begin{cases} x_{k,\boldsymbol{s}}(x_0, y_0) &= a_{k,\boldsymbol{s}} x_0 + b_{k,\boldsymbol{s}}, \\ y_{k,\boldsymbol{s}}(x_0, y_0) &= c_{k,\boldsymbol{s}} y_0 + d_{k,\boldsymbol{s}}. \end{cases} \quad k = 0, 1, \cdots, N-1. \tag{3.5}$$

The affine parameters in (3.5) are functions of itinerary $\boldsymbol{s}$. For a specific $\boldsymbol{s}$ they can be obtained in the following recursive way

$$\begin{cases} a_{k+1,\boldsymbol{s}} &= -2s_k a_{k,\boldsymbol{s}}, \\ b_{k+1,\boldsymbol{s}} &= 1 - 2s_k b_{k,\boldsymbol{s}}, \\ c_{k+1,\boldsymbol{s}} &= -\frac{1}{2} s_k c_{k,\boldsymbol{s}}, \\ d_{k+1,\boldsymbol{s}} &= \frac{1}{2} s_k (1 - d_{k,\boldsymbol{s}}), \end{cases} \qquad k = 0, \cdots, N-2, \tag{3.6}$$

with the starting point

$$\begin{cases} a_{0,\boldsymbol{s}} &= 1, \\ b_{0,\boldsymbol{s}} &= 0, \\ c_{0,\boldsymbol{s}} &= 1, \\ d_{0,\boldsymbol{s}} &= 0. \end{cases} \tag{3.7}$$

In fact the collection of 2^{N-1} itineraries one-to-one maps onto a partition[1] of the feasible space of x_0, i.e. the segment $[-1, +1]$. The itinerary $\mathbf{s}$ is a function of input x_0. For any specific itinerary $\mathbf{s}$, the admissible values of x_0 fall in a segment of length $1/2^{N-2}$, which is called a *cell* and denoted as $C_{\mathbf{s}} \triangleq [e_{l,\boldsymbol{s}}, e_{u,\boldsymbol{s}}]$. The two endpoints $e_{l,\boldsymbol{s}}$ and $e_{u,\boldsymbol{s}}$ of the cell associated with $\boldsymbol{s}$ are determined as

$$\begin{cases} e_{l,\boldsymbol{s}} &= \min\{\frac{-b_{N-1,\boldsymbol{s}}+1}{a_{N-1,\boldsymbol{s}}}, \frac{-b_{N-1,\boldsymbol{s}}-1}{a_{N-1,\boldsymbol{s}}}\}, \\ e_{u,\boldsymbol{s}} &= \max\{\frac{-b_{N-1,\boldsymbol{s}}+1}{a_{N-1,\boldsymbol{s}}}, \frac{-b_{N-1,\boldsymbol{s}}-1}{a_{N-1,\boldsymbol{s}}}\}. \end{cases} \tag{3.8}$$

This concept is illustrated in Fig.3.1. In the left part of Fig.1-(a), when $N = 2$, itinerary has one bit, i.e. $\mathbf{s} \in \{+1, -1\}$. The two corresponding cells are respectively the left and

[1]Here we ambiguously use the terminology partition, since every two adjacent cells overlap with their common endpoints. But this does not harm decoding procedure.

right half of the segment $[-1,+1]$. This concept is extended to length of $N=n+1$ in the right of Fig.1-(a), where $G^{(n)}(x)$ denotes n-fold composition of $G(\cdot)$. In fact, once the itinerary $\mathbf{s}_j$ is given, the endpoints $e_{l,\mathbf{s}_j}$ and $e_{u,\mathbf{s}_j}$ of the cell and the affine parameters $\{a_{k,\mathbf{s}_j}, b_{k,\mathbf{s}_j}, c_{k,\mathbf{s}_j}, d_{k,\mathbf{s}_j}\}$'s can all be determined as functions of $\mathbf{s}_j$, as shown in Fig.1-(b).

Next we discuss decoding schemes for the above baker's dynamic encoding system.

3.2.1 Maximum Likelihood(ML) Decoding

Under the AWGN channel assumption, the received signal can be represented as

$$\begin{cases} r_{x,n} = x_n + n_{x,n}, \\ r_{y,n} = y_n + n_{y,n}, \end{cases} \qquad n = 0, 1, \cdots, N-1, \tag{3.9}$$

where $n_{x,n},\ n_{y,n} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\sigma^2), n=0,1,\cdots,N-1$. We denote $\boldsymbol{r}_x = [r_{x,1}, r_{x,2}, \cdots, r_{x,N-1}]^T$ and $\boldsymbol{r}_y = [r_{y,1}, r_{y,2}, \cdots, r_{y,N-1}]^T$. The likelihood function of the observation sequences $\boldsymbol{r}_x, \boldsymbol{r}_y$ with given source (x_0, y_0) is

$$p(\boldsymbol{r}_x, \boldsymbol{r}_y | x_0, y_0) = (2\pi\sigma^2)^{-N} \exp\left\{ -\frac{\|\boldsymbol{r}_x - \boldsymbol{x}\|^2 + \|\boldsymbol{r}_y - \boldsymbol{y}\|^2}{2\sigma^2} \right\}. \tag{3.10}$$

The maximum likelihood estimate of source pair $\hat{x}_0^{ML}, \hat{y}_0^{ML}$ is

$$\begin{aligned} \{\hat{x}_0^{ML}, \hat{y}_0^{ML}\} &= \underset{-1\le x_0, y_0 \le 1}{\arg\max}\ p(\boldsymbol{r}_x, \boldsymbol{r}_y | x_0, y_0) \\ &= \underset{-1\le x_0, y_0 \le 1}{\arg\min} \sum_{k=0}^{N-1} \Big[\big(r_{x,k} - x_k(x_0,y_0)\big)^2 + \big(r_{y,k} - y_k(x_0,y_0)\big)^2 \Big]. \end{aligned} \tag{3.11}$$

The last equality emphasizes the fact that all x_k, y_k are all functions of x_0 and y_0.

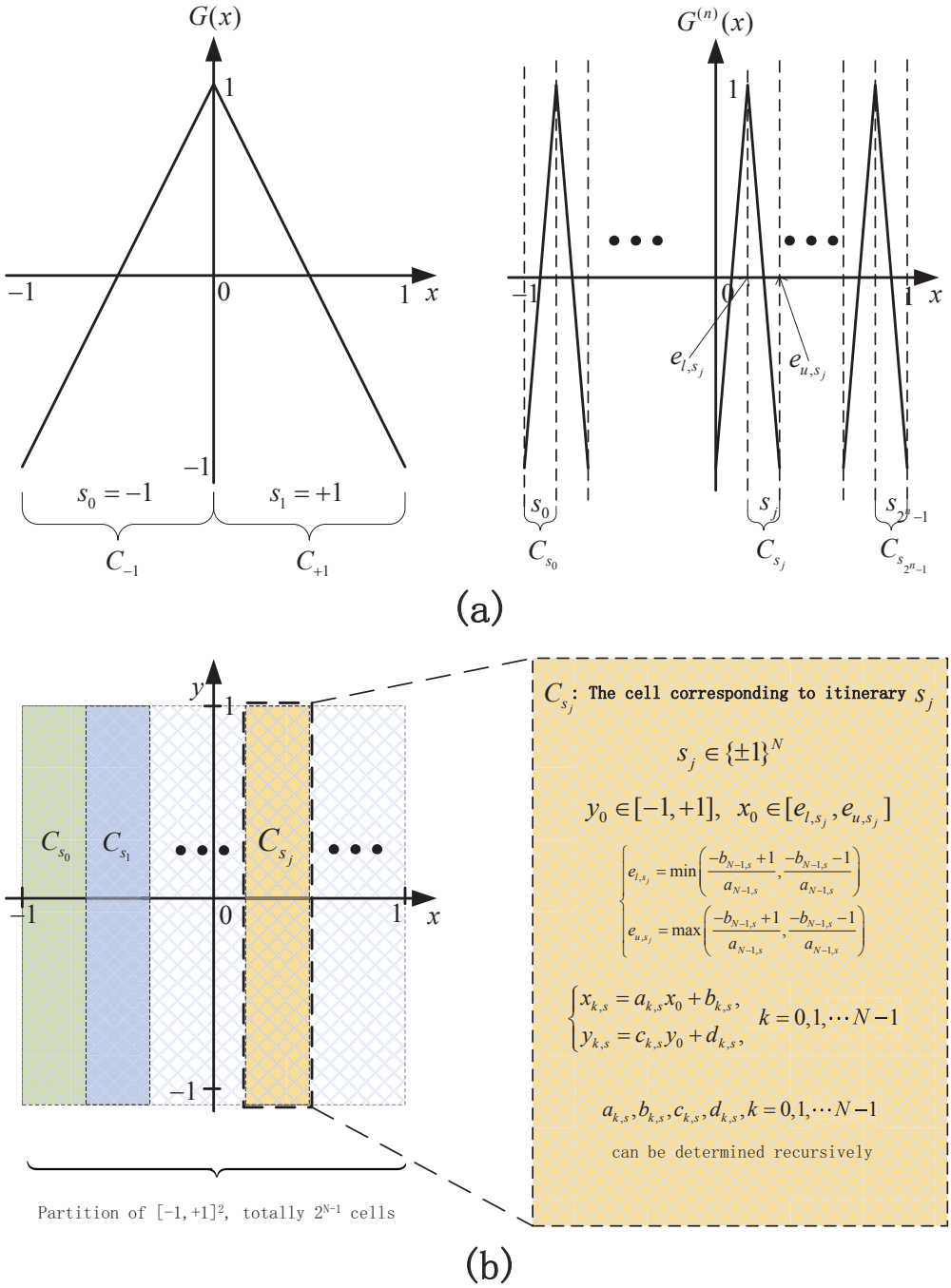

Figure 3.1: Partition and Itinerary— Left of (a): When $N = 1$, itinerary s has just one bit, $+1$ or -1; Right of (a): For general $N = n+1$, itinerary s has 2^n patterns. Each specific pattern $\mathbf{s}_j$ corresponds to one segment (cell) $C_{\mathbf{s}_j}$ of the feasible region; (b): The feasible region $[-1,+1]^2$ is partitioned into 2^{N-1} cells, with each cell $C_{\mathbf{s}_j}$ corresponding to one specific itinerary pattern $\mathbf{s}_j$. The parameters in the affine representation of the codewords and the endpoints of the cell can be determined once the itinerary $\mathbf{s}_j$ is given.

Based on connections between itineraries and cells discussed in (3.5)-(3.8), the original ML estimation problem in (3.11) can be further transformed into

$$\begin{aligned}\left(\hat{x}_0^{ML}, \hat{y}_0^{ML}\right) &= \arg\min_{\mathbf{s}, x_0 \in C_{\mathbf{s}}} \sum_{k=0}^{N-1} \left\{ \left[r_{x,k} - (a_{k,\boldsymbol{s}} x_0 + b_{k,\boldsymbol{s}})\right]^2 + \left[r_{y,k} - (c_{k,\boldsymbol{s}} y_0 + d_{k,\boldsymbol{s}})\right]^2 \right\} \\ &= \arg\min_{\boldsymbol{s}} \left\{ \min_{\substack{e_{1,\boldsymbol{s}} \le x_0 \le e_{2,\boldsymbol{s}} \\ -1 \le y_0 \le 1}} \sum_{k=0}^{N-1} \left\{ \left[r_{x,k} - (a_{k,\boldsymbol{s}} x_0 + b_{k,\boldsymbol{s}})\right]^2 + \left[r_{y,k} - (c_{k,\boldsymbol{s}} y_0 + d_{k,\boldsymbol{s}})\right]^2 \right\} \right\}. \end{aligned} \tag{3.12}$$

For any given itinerary $\boldsymbol{s}$, the inner-minimization problem in equation (3.12) is convex and quadratic. Without considering the constraints, its optimal solution $(x_{0,\boldsymbol{s}}^*, y_{0,\boldsymbol{s}}^*)$ is given in a closed form

$$\begin{cases} x_{0,\boldsymbol{s}}^* &= \frac{\boldsymbol{a}_{\boldsymbol{s}}^T (\boldsymbol{r}_{\boldsymbol{x}} - \boldsymbol{b}_{\boldsymbol{s}})}{\boldsymbol{a}_{\boldsymbol{s}}^T \boldsymbol{a}_{\boldsymbol{s}}}, \\ y_{0,\boldsymbol{s}}^* &= \frac{\boldsymbol{c}_{\boldsymbol{s}}^T (\boldsymbol{r}_{\boldsymbol{y}} - \boldsymbol{d}_{\boldsymbol{s}})}{\boldsymbol{c}_{\boldsymbol{s}}^T \boldsymbol{c}_{\boldsymbol{s}}}, \end{cases} \tag{3.13}$$

where, $\boldsymbol{a}_{\boldsymbol{s}} = \left[a_{0,\boldsymbol{s}}, \cdots, a_{N-1,\boldsymbol{s}}\right]^T$, $\boldsymbol{b}_{\boldsymbol{s}} = \left[b_{0,\boldsymbol{s}}, \cdots, b_{N-1,\boldsymbol{s}}\right]^T$, $\boldsymbol{c}_{\boldsymbol{s}} = \left[c_{0,\boldsymbol{s}}, \cdots, c_{N-1,\boldsymbol{s}}\right]^T$ and $\boldsymbol{d}_{\boldsymbol{s}} = \left[d_{0,\boldsymbol{s}}, \cdots, d_{N-1,\boldsymbol{s}}\right]^T$. Taking into account that the feasible (x_0, y_0) associated with $\boldsymbol{s}$ should lie within admissible range, a limiting procedure must be performed to obtain solution to the inner minimization with specific $\boldsymbol{s}$, i.e.

$$x_{0,\boldsymbol{s}}^{inner} = \begin{cases} e_{l,\boldsymbol{s}}, & \text{if } x_{0,\boldsymbol{s}}^* < e_{l,\boldsymbol{s}} \\ e_{u,\boldsymbol{s}}, & \text{if } x_{0,\boldsymbol{s}}^* > e_{u,\boldsymbol{s}} \\ x_{0,\boldsymbol{s}}^*, & \text{otherwise.} \end{cases}, \quad y_{0,\boldsymbol{s}}^{inner} = \begin{cases} -1, & \text{if } y_{0,\boldsymbol{s}}^* < -1 \\ +1, & \text{if } x_{0,\boldsymbol{s}}^* > +1 \\ y_{0,\boldsymbol{s}}^*, & \text{otherwise.} \end{cases}. \tag{3.14}$$

Since there are totally finite number of possible itinerary patterns, by enumerating all possible itineraries and selecting the $\{x_{0,\boldsymbol{s}}^{inner}, y_{0,\boldsymbol{s}}^{inner}\}$ which minimizes the outer mini-

mization, the ML estimation of (x_0, y_0) is obtained as

$$\left(\hat{x}_0^{ML}, \hat{y}_0^{ML}\right)=\arg\min_{\boldsymbol{s}}\left\{\sum_{k=0}^{N-1}\left\{\left[r_{x,k}-(a_{k,\boldsymbol{s}}x_{0,\boldsymbol{s}}^{inner}+b_{k,\boldsymbol{s}})\right]^2+\left[r_{y,k}-(c_{k,\boldsymbol{s}}y_{0,\boldsymbol{s}}^{inner}+d_{k,\boldsymbol{s}})\right]^2\right\}\right\}$$

The ML decoding scheme does not require *a priori* knowledge of the source's distribution. So it is applicable regardless of the probability distribution of the source.

3.2.2 Minimum Mean Square Error(MMSE) Decoding

The ML decoding method is not optimal in the sense of mean square error performance. In this subsection we focus on the MMSE solution to the baker's dynamic system. The MMSE estimator is given in a general form as [33]

$$\hat{X}^{MMSE}(y) = \mathsf{E}\{X|y\} = \int x f(x|y)\mathrm{d}x, \tag{3.15}$$

where X is random parameter to be determined and y is a specific realization of the noisy observation Y. It is worth noting that the above general solution usually cannot result in a closed form solution for concrete problems. Fortunately, under the uniform distribution assumption of the source signal, closed form MMSE estimator for baker's map can be obtained.

To provide the result of MMSE decoder, here we introduce the following notations

$$\begin{aligned} &A_1 = \|\boldsymbol{a}_{\boldsymbol{s}}\|^2; \quad B_1 = \boldsymbol{a}_{\boldsymbol{s}}^T(\boldsymbol{b}_{\boldsymbol{s}} - \boldsymbol{r}_x); \quad C_1 = \|\boldsymbol{b}_{\boldsymbol{s}} - \boldsymbol{r}_x\|^2; \\ &A_2 = \|\boldsymbol{c}_{\boldsymbol{s}}\|^2; \quad B_2 = \boldsymbol{c}_{\boldsymbol{s}}^T(\boldsymbol{d}_{\boldsymbol{s}} - \boldsymbol{r}_y); \quad C_2 = \|\boldsymbol{d}_{\boldsymbol{s}} - \boldsymbol{r}_y\|^2; \end{aligned} \tag{3.16}$$

and

$$E_1 = \exp\left\{\frac{B_1^2 - A_1C_1}{2\sigma^2 A_1}\right\};\;\; D_1 = Q\bigg(\frac{\sqrt{A_1}}{\sigma}e_{l,\boldsymbol{s}} + \frac{B_1}{\sigma\sqrt{A_1}}\bigg) - Q\bigg(\frac{\sqrt{A_1}}{\sigma}e_{u,\boldsymbol{s}} + \frac{B_1}{\sigma\sqrt{A_1}}\bigg);$$
$$E_2 = \exp\left\{\frac{B_2^2 - A_2C_2}{2\sigma^2 A_2}\right\};\;\; D_2 = Q\bigg(-\frac{\sqrt{A_2}}{\sigma} + \frac{B_2}{\sigma\sqrt{A_2}}\bigg) - Q\bigg(\frac{\sqrt{A_2}}{\sigma} + \frac{B_2}{\sigma\sqrt{A_2}}\bigg);$$
$$J_1 = \exp\left\{-\frac{1}{2\sigma^2}A_1\bigg(e_{l,\boldsymbol{s}} + \frac{B_1}{A_1}\bigg)^2\right\} - \exp\left\{-\frac{1}{2\sigma^2}A_1\bigg(e_{u,\boldsymbol{s}} + \frac{B_1}{A_1}\bigg)^2\right\};$$
$$J_2 = \exp\left\{-\frac{1}{2\sigma^2}A_2\bigg(1 + \frac{B_2}{A_2}\bigg)^2\right\} - \exp\left\{-\frac{1}{2\sigma^2}A_2\bigg(-1 + \frac{B_2}{A_2}\bigg)^2\right\}, \tag{3.17}$$

where the function $Q(\cdot)$ is the well known Gaussian-Q function which is defined as

$$Q(x) \triangleq \int_x^\infty \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}}\mathrm{d}t. \tag{3.18}$$

The MMSE estimator of x_0 and y_0 are given in a closed form as follows:

$$\hat{x}_0^{MMSE} = \frac{\sum_{\boldsymbol{s}} \sqrt{\frac{2\pi}{A_2}}E_1E_2D_2\bigg(\frac{\sigma}{A_1}J_1 - \frac{\sqrt{2\pi}B_1}{A_1^{3/2}}D_1\bigg)}{\sum_{\boldsymbol{s}} \frac{2\pi}{\sqrt{A_1A_2}}E_1E_2D_1D_2}, \tag{3.19}$$

$$\hat{y}_0^{MMSE} = \frac{\sum_{\boldsymbol{s}} \sqrt{\frac{2\pi}{A_1}}E_1E_2D_1\bigg(\frac{\sigma}{A_2}J_2 - \frac{\sqrt{2\pi}B_2}{A_2^{3/2}}D_2\bigg)}{\sum_{\boldsymbol{s}} \frac{2\pi}{\sqrt{A_1A_2}}E_1E_2D_1D_2}. \tag{3.20}$$

The detailed proof of the above result is rather involved and relegated to the appendix.

3.2.3 Mixed ML-MMSE Decoding Scheme

The MMSE estimator involves highly nonlinear numerical evaluations, like the Q-function, which are computation demanding and costly for implementation. In the next,

we introduce some kind of mixed ML and MMSE estimator for baker's analog code.

As previously discussed, once the itinerary is given, the analog codewords can be written as an affine function in (x_0, y_0). For specific itinerary $\boldsymbol{s}$, by packing the codewords $\boldsymbol{x}$ and $\boldsymbol{y}$ into one vector $\boldsymbol{v}$ and using (3.5), we can rewrite the baker's dynamic system as follows:

$$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \underbrace{\begin{bmatrix} \boldsymbol{a_s} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{c_s} \end{bmatrix}}_{\mathbf{G_s^T}} \underbrace{\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}}_{\mathbf{u}} + \underbrace{\begin{bmatrix} \boldsymbol{b_s} \\ \boldsymbol{d_s} \end{bmatrix}}_{\mathbf{t_s}} = \boldsymbol{G_s^T u} + \boldsymbol{t_s}, \tag{3.21}$$

where parameters $\boldsymbol{a_s}$, $\boldsymbol{b_s}$, $\boldsymbol{c_s}$ and $\boldsymbol{d_s}$ are defined in (3.6). Recall that in ML decoding a detection of the itinerary $\boldsymbol{s}$ can be obtained. By substituting $\boldsymbol{s}$ in (3.21) with the ML detection $\hat{\boldsymbol{s}}^{ML}$ and packing the received signals $\boldsymbol{r}_x$ and $\boldsymbol{r}_y$ into one vector $\boldsymbol{r} = [\boldsymbol{r}_x^T, \boldsymbol{r}_y^T]^T$, (3.9) can be expressed in a compact form

$$\boldsymbol{r}'_{\hat{\boldsymbol{s}}^{ML}} = \boldsymbol{r} - \boldsymbol{t}_{\hat{\boldsymbol{s}}^{ML}} = \boldsymbol{G}^T_{\hat{\boldsymbol{s}}^{ML}} \boldsymbol{u}. \tag{3.22}$$

Thus the baker's map code is equivalent to a $(2N, 2)$ linear analog code with encoder $\boldsymbol{G_s}$.

Now the problem to determine the source signal $\mathbf{u}$ in the above equation becomes the standard minimum MSE receiving problem, whose solution is the well known Wiener filter and given as [34]

$$\hat{\boldsymbol{u}}_{MMSE}(\hat{\boldsymbol{s}}^{ML}) = (\boldsymbol{G}_{\hat{\boldsymbol{s}}^{ML}} \boldsymbol{G}^T_{\hat{\boldsymbol{s}}^{ML}} + 3\sigma^2 \boldsymbol{I})^{-1} \boldsymbol{G}_{\hat{\boldsymbol{s}}^{ML}} \boldsymbol{r}'_{\hat{\boldsymbol{s}}^{ML}}. \tag{3.23}$$

A slicing operation then follows the above Wiener filtering to ensure the final estimate

$\hat{x}_0^{ML-MMSE}$ and $\hat{y}_0^{ML-MMSE}$ lie in $[e_{l,\hat{\mathbf{s}}^{ML}}, e_{u,\hat{\mathbf{s}}^{ML}}]$ and $[-1, +1]$ respectively.

For the mixed ML-MMSE method, ML decoding is performed to obtain $\hat{\boldsymbol{s}}^{ML}$. Then Wiener filtering and limiting procedure follows. The mixed ML-MMSE decoding method requires *a priori* knowledge of source and involves only linear computation operations.

3.2.4 Performance Analysis

In Fig.3.2 the different decoding algorithms' performance for the baker's analog codes with different length are plotted. E_u means the average power for each source signal and N_0 denotes the unilateral power spectral density, i.e. $N_0 = 2\sigma^2$. The ML, MMSE and ML-MMSE decoding algorithms have identical MSE performance for high SNR. In low SNR range, MMSE decoding method has best performance.

In the following, we analyze the MSE performance of the baker's dynamic coding system by considering the Cramer-Rao lower bound. CRLB is a lower bound for unbiased estimator [32]. It should be pointed out that, the ML decoding methods discussed above are biased estimator due to the slicing operations. However when SNR is large, the decoding error is sufficiently small such that the slicing rarely impact the decoding result. So CRLB can precisely predict the decoding error when SNR is modestly large and is useful a tool to understand the system's performance. This will also be verified by the following numerical results.

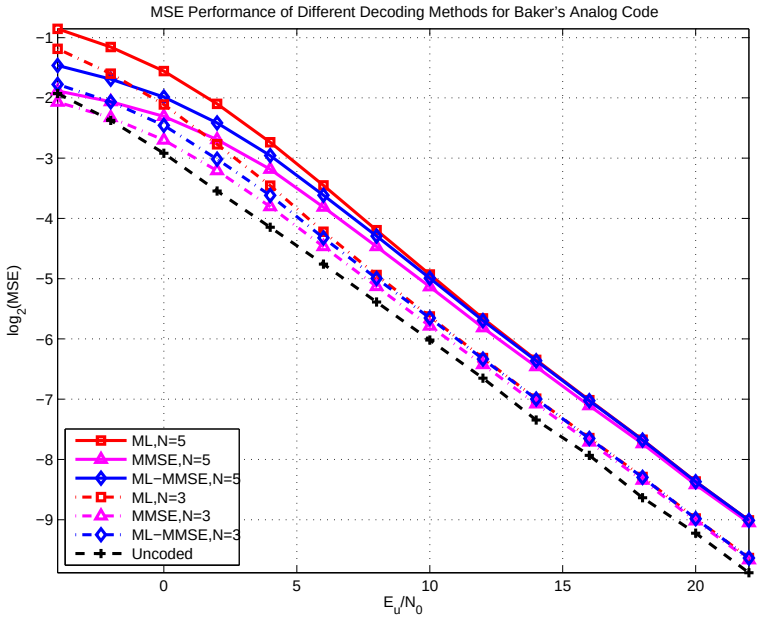

Figure 3.2: MSE performance of different decoding algorithms for baker's dynamic system

The Cramer-Rao lower bound for x_0 is given as [32]

$$CRLB_{x_0}^{baker} = -\mathsf{E}_{x_0}^{-1}\left\{\frac{\partial^2}{\partial x_0^2}\log p(\mathbf{r}_x, \mathbf{r}_y|x_0, y_0)\right\} \tag{3.24}$$

$$= -\mathsf{E}_{x_0}^{-1}\left\{\frac{\partial^2}{\partial x_0^2}\left(\frac{-1}{2\sigma^2}\sum_{k=0}^{N-1}\left((r_{x,k}-a_{k,\mathbf{s}}x_0-b_{k,\mathbf{s}})^2+(r_{y,k}-c_{k,\mathbf{s}}y_0-d_{k,\mathbf{s}})^2\right)\right)\right\} \tag{3.25}$$

$$= \frac{\sigma^2}{\sum_{k=0}^{N-1} a_{k,\mathbf{s}}^2} = \frac{3\sigma^2}{4^N-1}. \tag{3.26}$$

where $p(\mathbf{r}_x, \mathbf{r}_y|x_0, y_0)$ is defined in (3.10), and $\mathsf{E}_{x_0}(\cdot)$ denotes the expectation with respect to x_0. The recursive relations in (3.6) and the fact $s_k^2 = 1$ are used to obtain (3.26). Similarly the CRLB for y_0 obtained as

$$CRLB_{y_0}^{baker} = -E_{y_0}^{-1}\left\{\left(\frac{\partial}{\partial y_0}\log p(\mathbf{r}_x, \mathbf{r}_y|x_0, y_0)\right)^2\right\} = \frac{\sigma^2}{\sum_{k=0}^{N-1} c_k^2} = \frac{3\sigma^2}{4(1-(1/4)^N)}. \tag{3.27}$$

When N is modestly large, $CRLB_{x_0} \approx 3\sigma^2/4^N$. Each increment in N can decrease the decoding distortion of x_0 by $3/4$. Comparatively, increment in N improves little in y_0 determination, which is nearly a constant as $3\sigma^2/4$. The CRLB's reveals that the two sources are under unequal protection and there is insufficient coding gain on y_0. Recall that the x-sequence in codewords is obtained by continuously stretching and shifting the signal. Intuitively the signal is locally magnified. In comparison the y-sequence is obtained by compressing the signal. That's why the terms of 2^N and 2^{-N} appear in the denominator of CRLB for x_0 and y_0 respectively. This insight is verified by Fig.3.3, where separate MSE decoding performance of x_0 and y_0 are plotted with their CRLB's illustrated as benchmarks. Although x_0 has an obvious coding gain, y_0 is poorly protected and its distortion dominates the overall decoding performance.

From the CRLB analysis, we realize that the bottleneck of the baker's analog code lies in the weak protection to y_0. Thus to improve the baker's map code, effective protection should also be performed to y_0.

3.3 Improvement I—Mirrored Baker's Analog Code

As analyzed in the last section, the unsatisfying performance of the original baker's map lies in the poor protection of y_0. To enhance the protection of y_0, a natural idea is to perform a second original baker's map encoding by switching the roles of x_0 and y_0. Thus both x_0 and y_0 obtain balanced and effective protection. This idea leads to the improvement scheme to be discussed in this section—the mirrored baker's dynamic coding system. The mirrored baker's structure comprises two branches, with its first branch being the original baker's encoder and the second branch exchanging the roles

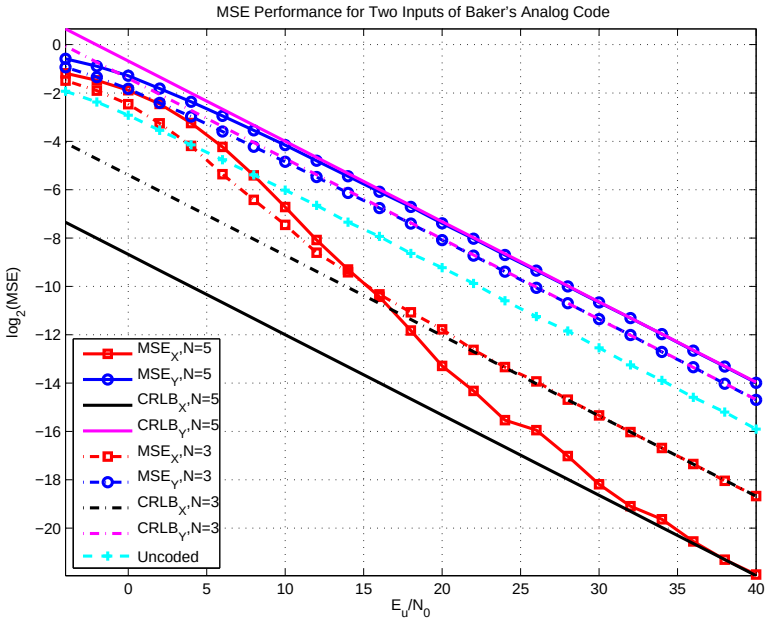

Figure 3.3: MSE performance for x_0 and y_0 of baker's system

of x_0 and y_0 to perform the original baker's encoding for a second time. For a given N, the mirrored baker's system forms a $(4N, 2)$ analog code.

Here we adjust our notations for the new system to make our following discussions clear. The two codewords associated with two branches are labeled with subscript 1 and 2 respectively. In the 1-st branch x_0 is tent map encoded and so does y_0 in the 2-nd branch. The codewords associated with x_0 and y_0 of the two branches are denoted as $\{\boldsymbol{x}_1, \boldsymbol{y}_1\}$ and $\{\boldsymbol{x}_2, \boldsymbol{y}_2\}$ respectively with their corresponding noisy observations as $\{\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}\}$ and $\{\boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y}\}$ respectively. The encoding procedure is expressed as

$$\begin{bmatrix} x_{1,n+1} \\ y_{1,n+1} \end{bmatrix} = F(x_{1,n}, y_{1,n}), \begin{bmatrix} y_{2,n+1} \\ x_{2,n+1} \end{bmatrix} = F(y_{2,n}, x_{2,n}),\ n = 0, \cdots, N-2; \qquad (3.28)$$

with $x_{1,0} = x_{2,0} = x_0$ and $y_{1,0} = y_{2,0} = y_0$. The observations are represented as:

$$\begin{cases} r_{j,x,n} = x_{j,n} + n_{j,x,n}, \\ r_{j,y,n} = y_{j,n} + n_{j,y,n}, \end{cases}, \ j = 1, 2; \ n = 0, 1, \cdots, N-1. \tag{3.29}$$

The mirrored baker's dynamic system has two itineraries $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ from the 1-st and 2-nd branch respectively, the two of which compose the entire itinerary for the mirrored baker's system. As previously discussed, $\mathbf{s}_1$ indicates a partition of the feasible domain of x_0. So does $\mathbf{s}_2$ to y_0. The entire feasible domain for the source pair (x_0, y_0), which is a 2×2 square centered at the origin on the plane, is uniformly divided into $2^{(2N-2)}$ cells, with each cell being a tiny square having edge of length $2^{-(N-2)}$. Assuming that the source (x_0, y_0) is known to live in some specific cell, the itineraries $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ can be determined and the codewords can be expressed as affine functions:

$$\begin{cases} x_{1,k,\boldsymbol{s}_1}(x_0, y_0) = a_{1,k,\boldsymbol{s}_1} x_0 + b_{1,k,\boldsymbol{s}_1}, \\ y_{1,k,\boldsymbol{s}_1}(x_0, y_0) = c_{1,k,\boldsymbol{s}_1} y_0 + d_{1,k,\boldsymbol{s}_1}, \end{cases} \begin{cases} x_{2,k,\boldsymbol{s}_2}(x_0, y_0) = a_{2,k,\boldsymbol{s}_2} x_0 + b_{2,k,\boldsymbol{s}_2}, \\ y_{2,k,\boldsymbol{s}_2}(x_0, y_0) = c_{2,k,\boldsymbol{s}_2} y_0 + d_{2,k,\boldsymbol{s}_2}, \end{cases} \tag{3.30}$$

with $k = 0, 1, \cdots, N-2$. The parameters $\{a_{1,k,\boldsymbol{s}_1}, b_{1,k,\boldsymbol{s}_1}, c_{1,k,\boldsymbol{s}_1}, d_{1,k,\boldsymbol{s}_1}\}$ and $\{a_{2,k,\boldsymbol{s}_2}, b_{2,k,\boldsymbol{s}_2}, c_{2,k,\boldsymbol{s}_2}, d_{2,k,\boldsymbol{s}_2}\}$ are for the 1-st and the 2-nd branch respectively and can be determined recursively for $k = 0, \cdots, N-2$ as follows

$$\begin{cases} a_{1,k+1,\boldsymbol{s}_1} = -2 s_{1,k} a_{1,k,\boldsymbol{s}_1}, \\ b_{1,k+1,\boldsymbol{s}_1} = 1 - 2 s_{1,k} b_{1,k,\boldsymbol{s}_1}, \\ c_{1,k+1,\boldsymbol{s}_1} = -\frac{1}{2} s_{1,k} c_{1,k,\boldsymbol{s}_1}, \\ d_{1,k+1,\boldsymbol{s}_1} = \frac{1}{2} s_{1,k} (1 - d_{1,k,\boldsymbol{s}_1}), \end{cases} \begin{cases} c_{2,k+1,\boldsymbol{s}_2} = -2 s_{2,k} c_{2,k,\boldsymbol{s}_2}, \\ d_{2,k+1,\boldsymbol{s}_2} = 1 - 2 s_{2,k} d_{2,k,\boldsymbol{s}_2}, \\ a_{2,k+1,\boldsymbol{s}_2} = -\frac{1}{2} s_{2,k} a_{2,k,\boldsymbol{s}_2}, \\ b_{2,k+1,\boldsymbol{s}_2} = \frac{1}{2} s_{2,k} (1 - b_{2,k,\boldsymbol{s}_2}), \end{cases} \tag{3.31}$$

with the starting point

$$\begin{cases} a_{1,0,\boldsymbol{s}_1} = a_{2,0,\boldsymbol{s}_2} = 1, \\ b_{1,0,\boldsymbol{s}_1} = b_{2,0,\boldsymbol{s}_2} = 0, \\ c_{1,0,\boldsymbol{s}_1} = c_{2,0,\boldsymbol{s}_2} = 1, \\ d_{1,0,\boldsymbol{s}_1} = d_{2,0,\boldsymbol{s}_2} = 0. \end{cases} \tag{3.32}$$

We denote $\boldsymbol{a}_{j,\boldsymbol{s}_j} = [a_{j,0,\boldsymbol{s}_j}, a_{j,1,\boldsymbol{s}_j}, \cdots, a_{j,N-1,\boldsymbol{s}_j}]^T$, $j = 1, 2$ and define $\boldsymbol{b}_{j,\boldsymbol{s}_j}$, $\boldsymbol{c}_{j,\boldsymbol{s}_j}$ and $\boldsymbol{d}_{j,\boldsymbol{s}_j}$ in the same way for $j = 1, 2$.

For a specific itinerary $\{\boldsymbol{s}_1, \boldsymbol{s}_2\}$, we denote its indicated admissible cell has projection $C_{\mathbf{s}_1}$ onto x_0 feasible domain and projection $C_{\mathbf{s}_2}$ onto y_0 feasible domain, i.e.

$$x_0 \in C_{\mathbf{s}_1} = [e_{1,l,\boldsymbol{s}_1}, e_{1,u,\boldsymbol{s}_1}], \;\; y_0 \in C_{\mathbf{s}_2} = [e_{2,l,\boldsymbol{s}_2}, e_{2,u,\boldsymbol{s}_2}], \text{with} \tag{3.33}$$

$$\begin{cases} e_{1,l,\boldsymbol{s}_1} = \min\{\frac{-b_{1,N-1,\boldsymbol{s}_1}+1}{a_{1,N-1,\boldsymbol{s}_1}}, \frac{-b_{1,N-1,\boldsymbol{s}_1}-1}{a_{N-1,\boldsymbol{s}_1}}\}, \\ e_{1,u,\boldsymbol{s}_1} = \max\{\frac{-b_{1,N-1,\boldsymbol{s}_1}+1}{a_{1,N-1,\boldsymbol{s}_1}}, \frac{-b_{1,N-1,\boldsymbol{s}_1}-1}{a_{N-1,\boldsymbol{s}_1}}\}, \end{cases} \begin{cases} e_{2,l,\boldsymbol{s}_2} = \min\{\frac{-d_{2,N-1,\boldsymbol{s}_2}+1}{c_{2,N-1,\boldsymbol{s}_2}}, \frac{-d_{2,N-1,\boldsymbol{s}_2}-1}{c_{2,N-1,\boldsymbol{s}_2}}\}, \\ e_{2,u,\boldsymbol{s}_2} = \max\{\frac{-d_{2,N-1,\boldsymbol{s}_2}+1}{c_{2,N-1,\boldsymbol{s}_2}}, \frac{-d_{2,N-1,\boldsymbol{s}_2}-1}{c_{2,N-1,\boldsymbol{s}_2}}\}, \end{cases}$$

Next we discuss decoding methods for the mirrored baker's dynamic system. These decoding methods are obtained by straightforwardly extending the results for the original baker's system. In the following main results are provided with details omitted.

3.3.1 ML Decoding

In this subsection, the ML decoding of the mirrored baker's map code is presented. The estimate $\hat{x}_0^{ML}, \hat{y}_0^{ML}$ maximizing the likelihood function is equivalently given as

$$\left(\hat{x}_0^{ML}, \hat{y}_0^{ML}\right) = \underset{\boldsymbol{s}_1, \boldsymbol{s}_2}{\arg\min} \left\{ \min_{\substack{e_{1,l,\boldsymbol{s}_1} \le x_0 \le e_{1,u,\boldsymbol{s}_1} \\ e_{2,l,\boldsymbol{s}_2} \le y_0 \le e_{2,u,\boldsymbol{s}_2}}} \sum_{j=1}^{2} \sum_{k=0}^{N-1} \left\{ \left[r_{j,x,k} - (a_{j,k,\boldsymbol{s}_j} x_0 + b_{j,k,\boldsymbol{s}_j}) \right]^2 + \left[r_{j,y,k} - (c_{j,k,\boldsymbol{s}_j} y_0 + d_{j,k,\boldsymbol{s}_j}) \right]^2 \right\} \right\}. \quad (3.34)$$

For a given pair of sequences $\{\boldsymbol{s}_1, \boldsymbol{s}_2\}$, the optimal solution of the inner minimization of the above equation is given as

$$\begin{cases} x^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} &= \frac{\boldsymbol{a}_{\boldsymbol{s}_1}^T (\boldsymbol{r}_{1,x} - \boldsymbol{b}_{\boldsymbol{s}_1}) + \boldsymbol{a}_{\boldsymbol{s}_2}^T (\boldsymbol{r}_{2,x} - \boldsymbol{b}_{\boldsymbol{s}_2})}{\boldsymbol{a}_{\boldsymbol{s}_1}^T \boldsymbol{a}_{\boldsymbol{s}_1} + \boldsymbol{a}_{\boldsymbol{s}_2}^T \boldsymbol{a}_{\boldsymbol{s}_2}}, \\ y^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} &= \frac{\boldsymbol{c}_{\boldsymbol{s}_1}^T (\boldsymbol{r}_{1,y} - \boldsymbol{d}_{\boldsymbol{s}_1}) + \boldsymbol{c}_{\boldsymbol{s}_2}^T (\boldsymbol{r}_{2,y} - \boldsymbol{d}_{\boldsymbol{s}_2})}{\boldsymbol{c}_{\boldsymbol{s}_1}^T \boldsymbol{c}_{\boldsymbol{s}_1} + \boldsymbol{c}_{\boldsymbol{s}_2}^T \boldsymbol{c}_{\boldsymbol{s}_2}}, \end{cases} \quad (3.35)$$

followed by the hard limiter:

$$x^{inner}_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} = \begin{cases} e_{1,l,\boldsymbol{s}_1}, \text{if } x^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} < e_{1,l,\boldsymbol{s}_1} \\ e_{1,u,\boldsymbol{s}_1}, \text{if } x^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} > e_{1,u,\boldsymbol{s}_1} \\ x^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2}, \text{otherwise.} \end{cases} \quad y^{inner}_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} = \begin{cases} e_{2,l,\boldsymbol{s}_2}, \text{if } y^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} < e_{2,l,\boldsymbol{s}_2} \\ e_{2,u,\boldsymbol{s}_2}, \text{if } y^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} > e_{2,u,\boldsymbol{s}_2} \\ y^*_{0,\boldsymbol{s}}, \text{otherwise.} \end{cases} . \quad (3.36)$$

The ML estimation is given by selecting the $(x^{inner}_{0,\boldsymbol{s}_1,\boldsymbol{s}_2}, y^{inner}_{0,\boldsymbol{s}_1,\boldsymbol{s}_2})$ among different itineraries $\{\boldsymbol{s}_1, \boldsymbol{s}_2\}$ which minimizes the outer minimization in (3.34).

3.3.2 MMSE Decoding

To introduce the MMSE decoding results for mirrored baker's system, we adopt the following notations:

$$\begin{cases} \bar{A}_1 = \|\boldsymbol{a}_{\boldsymbol{s}_1}\|^2 + \|\boldsymbol{a}_{\boldsymbol{s}_2}\|^2; \\ \bar{B}_1 = \boldsymbol{a}_{\boldsymbol{s}_1}^T(\boldsymbol{b}_{\boldsymbol{s}_1} - \boldsymbol{r}_{1,x}) + \boldsymbol{a}_{\boldsymbol{s}_2}^T(\boldsymbol{b}_{\boldsymbol{s}_2} - \boldsymbol{r}_{2,x}); \\ \bar{C}_1 = \|\boldsymbol{b}_{\boldsymbol{s}_1} - \boldsymbol{r}_{1,x}\|^2 + \|\boldsymbol{b}_{\boldsymbol{s}_2} - \boldsymbol{r}_{2,x}\|^2; \end{cases} \quad \begin{cases} \bar{A}_2 = \|\boldsymbol{c}_{\boldsymbol{s}_1}\|^2 + \|\boldsymbol{c}_{\boldsymbol{s}_2}\|^2; \\ \bar{B}_2 = \boldsymbol{c}_{\boldsymbol{s}_1}^T(\boldsymbol{d}_{\boldsymbol{s}_1} - \boldsymbol{r}_{1,y}) + \boldsymbol{c}_{\boldsymbol{s}_2}^T(\boldsymbol{d}_{\boldsymbol{s}_2} - \boldsymbol{r}_{2,y}); \\ \bar{C}_2 = \|\boldsymbol{d}_{\boldsymbol{s}_1} - \boldsymbol{r}_{1,y}\|^2 + \|\boldsymbol{d}_{\boldsymbol{s}_2} - \boldsymbol{r}_{2,y}\|^2; \end{cases} \tag{3.37}$$

$$\bar{E}_j = \exp\left\{\frac{\bar{B}_j^2 - \bar{A}_j\bar{C}_j}{2\sigma^2\bar{A}_j}\right\}; \bar{D}_j = Q\left(\frac{\sqrt{\bar{A}_j}}{\sigma}e_{j,l,\boldsymbol{s}_j} + \frac{\bar{B}_j}{\sigma\sqrt{\bar{A}_j}}\right) - Q\left(\frac{\sqrt{\bar{A}_j}}{\sigma}e_{j,u,\boldsymbol{s}_j} + \frac{\bar{B}_j}{\sigma\sqrt{\bar{A}_j}}\right);$$

$$\bar{J}_j = \exp\left\{-\frac{1}{2\sigma^2}\bar{A}_j\left(e_{j,l,\boldsymbol{s}_j} + \frac{\bar{B}_j}{\bar{A}_j}\right)^2\right\} - \exp\left\{-\frac{1}{2\sigma^2}\bar{A}_j\left(e_{j,u,\boldsymbol{s}_j} + \frac{\bar{B}_j}{\bar{A}_j}\right)^2\right\}, j = 1, 2.$$

The calculation of the MMSE estimation still follows similar lines as discussed for single baker system. The major difference is that since the sign sequence of y_0 contributes to the itinerary, the integration of y_0 should be decomposed into parts over different $C_{\boldsymbol{s}_2}$'s. The MMSE estimation of x_0 can be given as:

$$\begin{aligned} \hat{x}_0^{MMSE} &= E\{x_0|\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}, \boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y}\} = \int_{-1}^{+1} x_0 f(x_0|\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}, \boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y})\mathrm{d}x_0 \\ &= \sum_{\boldsymbol{s}_1\}} \int_{C_{\boldsymbol{s}_1}} x_0 \frac{f(\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}, \boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y}|x_0)f(x_0)}{f(\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}, \boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y})}\mathrm{d}x_0 \end{aligned} \tag{3.38}$$

$$\begin{aligned} &= \frac{1}{4f(\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}, \boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y})} \sum_{\{\boldsymbol{s}_1\}} \int_{C_{\boldsymbol{s}_1}} x_0 \sum_{\{\boldsymbol{s}_2\}} \int_{C_{\boldsymbol{s}_2}} f(\boldsymbol{r}_{1,x}, \boldsymbol{r}_{1,y}, \boldsymbol{r}_{2,x}, \boldsymbol{r}_{2,y}|x_0, y_0)\mathrm{d}y_0\mathrm{d}x_0 \\ &= \frac{\sum_{\{\boldsymbol{s}_1, \boldsymbol{s}_2\}} \sqrt{\frac{2\pi}{\bar{A}_2}}\bar{E}_1\bar{E}_2\bar{D}_2\left(\frac{\sigma}{\bar{A}_1}\bar{J}_1 - \frac{\sqrt{2\pi}\bar{B}_1}{\bar{A}_1^{3/2}}\bar{D}_1\right)}{\sum_{\{\boldsymbol{s}_1, \boldsymbol{s}_2\}} \frac{2\pi}{\sqrt{\bar{A}_1\bar{A}_2}}\bar{E}_1\bar{E}_2\bar{D}_1\bar{D}_2}. \end{aligned} \tag{3.39}$$

Similarly, the MMSE estimation of y_0 for the mirrored baker's map code is given as

follows

$$\hat{y}_0^{MMSE} = \frac{\sum\limits_{\{\boldsymbol{s}_1,\boldsymbol{s}_2\}} \sqrt{\frac{2\pi}{\bar{A}_1}} \bar{E}_1 \bar{E}_2 \bar{D}_1 \left(\frac{\sigma}{\bar{A}_2} \bar{J}_2 - \frac{\sqrt{2\pi}\bar{B}_2}{\bar{A}_2^{3/2}} \bar{D}_2 \right)}{\sum\limits_{\{\boldsymbol{s}_1,\boldsymbol{s}_2\}} \frac{2\pi}{\sqrt{\bar{A}_1 \bar{A}_2}} \bar{E}_1 \bar{E}_2 \bar{D}_1 \bar{D}_2}. \tag{3.40}$$

3.3.3 ML-MMSE Decoding

If the itinerary $\{\boldsymbol{s}_1, \boldsymbol{s}_2\}$ is given, the codewords of the mirrored baker's map system can be represented as an affine function of the original source (x_0, y_0). The the corresponding coefficients can be determined recursively by using equations (3.31) and (3.32). Thus the mirrored baker's dynamic system can be rewritten as the following:

$$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{y}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{\boldsymbol{s}_1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{c}_{\boldsymbol{s}_1} \\ \boldsymbol{a}_{\boldsymbol{s}_2} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{c}_{\boldsymbol{s}_2} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} \boldsymbol{b}_{\boldsymbol{s}_1} \\ \boldsymbol{d}_{\boldsymbol{s}_1} \\ \boldsymbol{b}_{\boldsymbol{s}_2} \\ \boldsymbol{d}_{\boldsymbol{s}_2} \end{bmatrix} = \boldsymbol{G}_{\boldsymbol{s}_1,\boldsymbol{s}_2}^T \boldsymbol{u} + \boldsymbol{t}_{\boldsymbol{s}_1,\boldsymbol{s}_2} \tag{3.41}$$

We can first perform the ML estimation discussed in previous subsection 3.3.1 and thus obtain the ML detection of the itinerary $\{\hat{s}_1^{ML}, \hat{s}_2^{ML}\}$. By taking the ML detection of the itinerary as true value, the linear MMSE estimator is invoked to estimate original value of $\{x_0, y_0\}$ as the follows:

$$\hat{\boldsymbol{u}}_{MMSE}(\hat{\boldsymbol{s}}_1^{ML}, \hat{\boldsymbol{s}}_2^{ML}) = \left(\boldsymbol{G}_{\hat{\boldsymbol{s}}_1^{ML},\hat{\boldsymbol{s}}_2^{ML}} \boldsymbol{G}_{\hat{\boldsymbol{s}}_1^{ML},\hat{\boldsymbol{s}}_2^{ML}}^T + 3\sigma^2 \boldsymbol{I} \right)^{-1} \boldsymbol{G}_{\hat{\boldsymbol{s}}^{ML}} \left(\boldsymbol{r} - \boldsymbol{t}_{\hat{\boldsymbol{s}}_1^{ML},\hat{\boldsymbol{s}}_2^{ML}} \right). \tag{3.42}$$

Then a limiting procedure is performed to obtain admissible decoding results.

3.4 Improvement II—Single-Input(1-D) Baker's Analog Code

Inspired by the performance analysis in section 3.2.4, to enhance the original baker's map performance, effective protection must be performed equally to all sources. Besides the mirrored structure proposed in last section, here we propose an alternative improving strategy is to feed the y-sequence with input x_0, which actually forms a single-input(1-D) baker's analog code. By feeding the two inputs of original baker's map with one source x_0, the problem of poor protection of y_0 vanishes and protection of x_0 is enhanced. In other words, the protection to all sources are equal and strengthened. Furthermore, another unconspicuous yet profound aspect of motivation of this 1-D scheme is that it performs a hidden repetition code of the itinerary, which is explained in full details as follows.

As pointed out in the papers [30] and [29], reliably determining the itinerary is a key factor impacting decoding performance. In the original baker's analog coding system, the y-sequence does not help to protect the itinerary since each of its signal is uncorrelated with x_0. Recall that the codeword of y-sequence of the baker's system is generated by inverse tent map function using sign sequence from the x-sequence. By feeding the y-sequence with x_0, we have $y_1 = G^{-1}_{\mathsf{sign}(x_0)}\big(x_0\big)$. Equivalently $x_0 = G\big(y_1\big)$. So actually y_1 can be regarded as the state immediately before x_0 in the tent dynamic system, which we denote as x_{-1}. Following this manner, we can regard y_i as the immediate previous state of y_{i-1} in a tent map dynamic sequence for $i = 2, \cdots, N-1$. Thus by rewriting the y-sequence signal as $\{y_{N-1}, y_{N-2}, \cdots, y_0\} \triangleq \{x_{-(N-1)}, x_{-(N-2)}, \cdots, x_0\}$ and concatenating it with the x-sequence signals, we actually obtain a long tent map analog

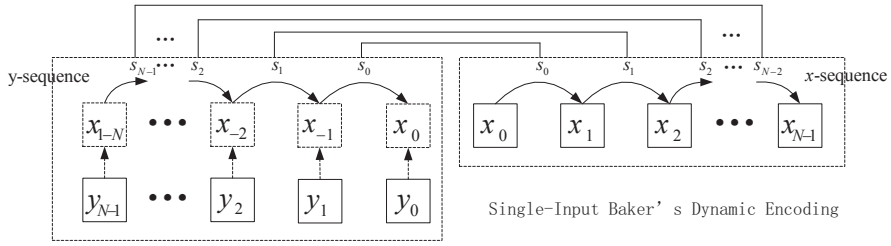

Figure 3.4: 1-D baker's dynamic encoding system

code (except that there are two copies of x_0 here). Moreover this obtained equivalent tent map sequence has its special pattern: the first half itinerary is reversely identical with the second half itinerary. In other words the 1-D baker's analog code actually constructs a hidden *repetition code* for the itinerary sequence. Both the x- and y- sequences now become analog "parity bits" of the itinerary. This interesting alternative view of the 1-D baker dynamic system is illustrated in Fig.3.4.

Next, sticking to the notations introduced above for the baker's system, we give out the decoding results for this one dimensional baker analog code.

3.4.1 ML Decoding Scheme

Similar to previous discussion, for each given itinerary $\boldsymbol{s}$, the optimal solution to inner minimization $x^*_{0,\boldsymbol{s}}$ is obtained by

$$x^{inner}_{0,\boldsymbol{s}} = \begin{cases} e_{l,\boldsymbol{s}}, \text{if } x^*_{0,\boldsymbol{s}} < e_{l,\boldsymbol{s}}, \\ e_{u,\boldsymbol{s}}, \text{if } x^*_{0,\boldsymbol{s}} > e_{u,\boldsymbol{s}}, \\ x^*_{0,\boldsymbol{s}_1,\boldsymbol{s}_2} \text{otherwise}, \end{cases} \quad \text{with } x^*_{0,\boldsymbol{s}} = \frac{\boldsymbol{a}_{\boldsymbol{s}}^T(\boldsymbol{r_x}-\boldsymbol{b_s})+\boldsymbol{c}_{\boldsymbol{s}}^T(\boldsymbol{r_y}-\boldsymbol{d_s})}{\boldsymbol{a}_{\boldsymbol{s}}^T\boldsymbol{a_s}+\boldsymbol{c}_{\boldsymbol{s}}^T\boldsymbol{c_s}}. \tag{3.43}$$

The ML estimate is obtained by going over all possible itineraries and selecting the $x_{0,\boldsymbol{s}}^{inner}$ which minimizes the likelihood function.

3.4.2 MMSE Decoding Scheme

Defining the following parameters

$$A = \|\boldsymbol{a_s}\|^2 + \|\boldsymbol{c_s}\|^2; \quad B = \boldsymbol{a}_{\boldsymbol{s}}^T(\boldsymbol{b_s} - \boldsymbol{r}_x) + \boldsymbol{c}_{\boldsymbol{s}}^T(\boldsymbol{d_s} - \boldsymbol{r}_y); \quad C = \|\boldsymbol{b_s} - \boldsymbol{r}_x\|^2 + \|\boldsymbol{d_s} - \boldsymbol{r}_y\|^2;$$
$$E = \exp\left\{\frac{B^2 - AC}{2\sigma^2 A}\right\}; \;\; D = Q\left(\frac{\sqrt{A}}{\sigma}e_{l,\boldsymbol{s}} + \frac{B}{\sigma\sqrt{A}}\right) - Q\left(\frac{\sqrt{A}}{\sigma}e_{u,\boldsymbol{s}} + \frac{B}{\sigma\sqrt{A}}\right);$$
$$J = \exp\left\{-\frac{1}{2\sigma^2}A\left(e_{l,\boldsymbol{s}} + \frac{B}{A}\right)^2\right\} - \exp\left\{-\frac{1}{2\sigma^2}A\left(e_{u,\boldsymbol{s}} + \frac{B}{A}\right)^2\right\}, \tag{3.44}$$

The MMSE estimate if given as:

$$\hat{x}_0^{MMSE} = \frac{\sum_{\boldsymbol{s}} \sqrt{\frac{2\pi}{A_2}} E\left(\frac{\sigma}{A}J - \frac{\sqrt{2\pi}B}{A^{3/2}}D\right)}{\sum_{\boldsymbol{s}} \sqrt{\frac{2\pi}{A}} ED}. \tag{3.45}$$

3.4.3 ML-MMSE Decoding Scheme

Assume that the ML detection of the itinerary is $\hat{\mathrm{s}}_{ML}$, then the received signal can be written in an affine form of x_0 as

$$\begin{bmatrix} \boldsymbol{r}_x \\ \boldsymbol{r}_y \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_{\hat{\boldsymbol{s}}^{ML}} \\ \boldsymbol{c}_{\hat{\boldsymbol{s}}^{ML}} \end{bmatrix} x_0 + \begin{bmatrix} \boldsymbol{b}_{\hat{\boldsymbol{s}}^{ML}} \\ \boldsymbol{d}_{\hat{\boldsymbol{s}}^{ML}} \end{bmatrix}. \tag{3.46}$$

The linear MMSE estimate is obtained by performing limiting procedure to the following value

$$\hat{x}_0^{MMSE}(\hat{\boldsymbol{s}}^{ML}) = \frac{\boldsymbol{a}_{\hat{\boldsymbol{s}}^{ML}}^T(\boldsymbol{r}_x - \boldsymbol{b}_{\hat{\boldsymbol{s}}^{ML}}) + \boldsymbol{c}_{\hat{\boldsymbol{s}}^{ML}}^T(\boldsymbol{r}_y - \boldsymbol{d}_{\hat{\boldsymbol{s}}^{ML}})}{\|\boldsymbol{a}_{\hat{\boldsymbol{s}}^{ML}}\|^2 + \|\boldsymbol{c}_{\hat{\boldsymbol{s}}^{ML}}\|^2 + 3\sigma^2}. \tag{3.47}$$

3.5 Simulation Results and Discussions

In this section numerical results and discussions are presented. The MSE performance of ML, MMSE and ML-MMSE decoding algorithms for mirrored baker's and single-input baker's system are presented in Fig.3.5 and Fig.3.6 respectively, where E_u represents the average power for each source signal and N_0 denotes the unilateral power spectral density. In our experiment the source signals are independent and uniformly distributed over $[-1, +1]$. For each coding system, codes with $N = 3$ and $N = 5$ are tested. The associated CRLB's(determined explicitly in equation (3.48)) and uncoded performance are plotted to serve as benchmarks. Numerical results verify the validity of the decoding algorithms developed in previous sections and show that both of the mirrored and single-input structure have improved MSE performance of the original baker's coding system.

Fig.3.7 compares the performance of the mirrored and single-input baker's map and the tent map analog codes proposed in [27,28], where coderate of $1/6$ and $1/10$ are considered for each coding scheme. Although tent map encoding scheme can be proved to have a lower CRLB, its actual performance is disadvantageous to the improved baker's schemes over a wide SNR range.

Generally the distortion of analog transmission systems can be decomposed into two

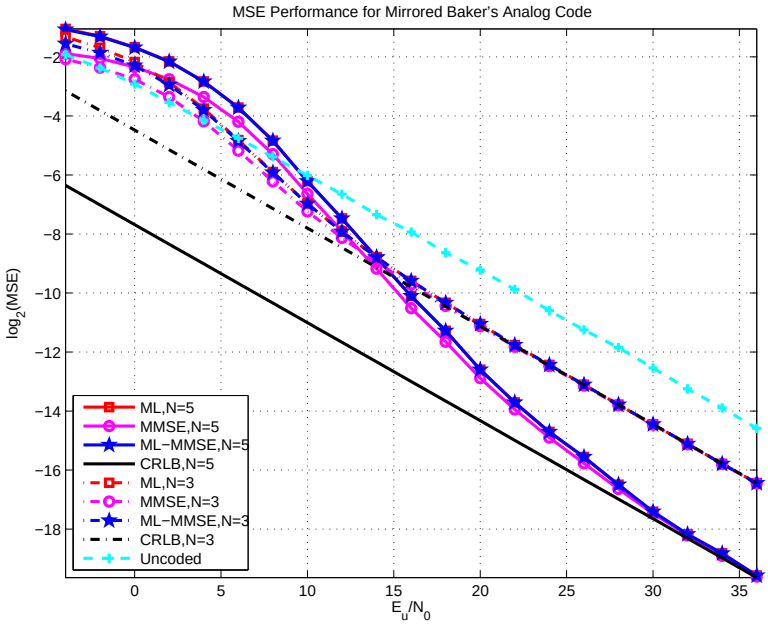

Figure 3.5: MSE of different decoding algorithms for mirrored baker's analog code

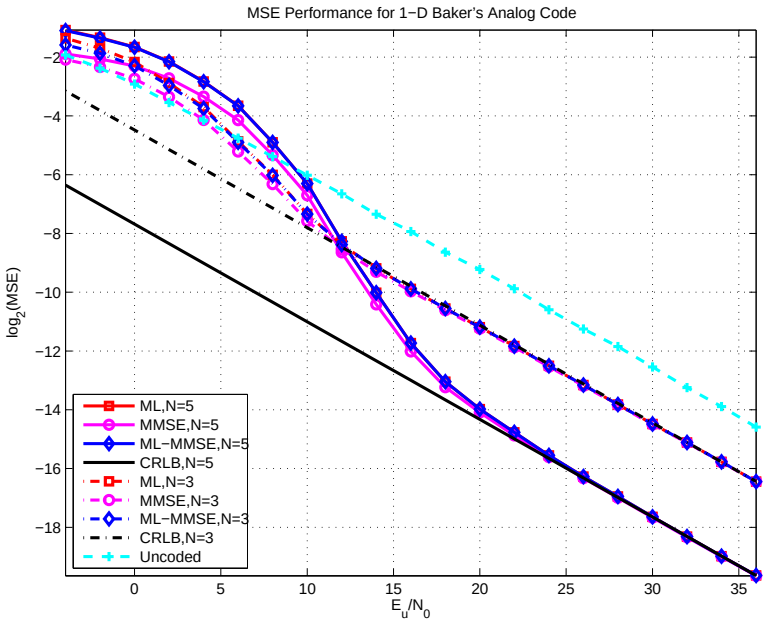

Figure 3.6: MSE performance of different decoding algorithms for single-input baker's analog code

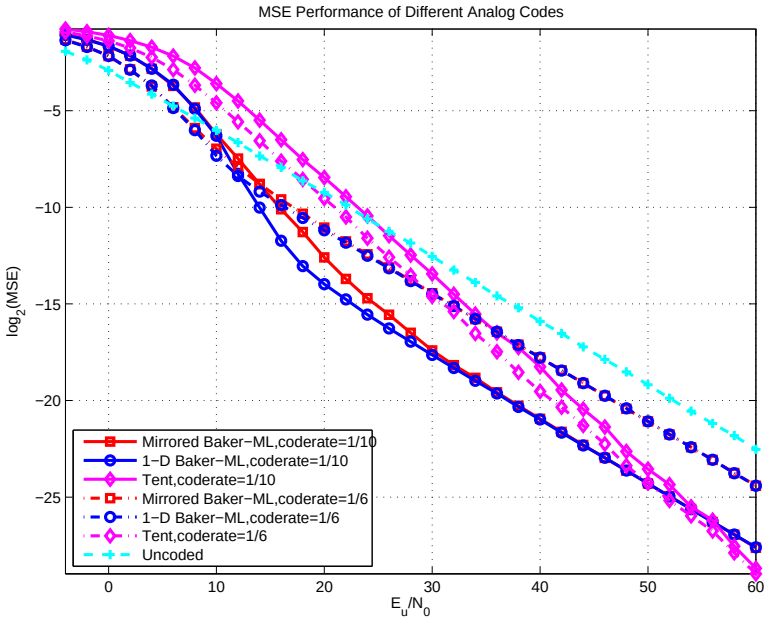

Figure 3.7: MSE performance of tent map code, mirrored baker's map code and single-input baker's map code(coderate 1/10)

parts [13]: anomalous distortion and the weak distortion. Weak distortion, stemming from the channel noise, can become very small and close to zero as long as the channel noise is sufficiently small. As analyzed in [26], to reduce the distortion of estimation, transmitted signal must be stretched as much as possible, which can be intuitively seen as "amplifying" the signal. However due to transmission power constraint, transmitted signals have to be bounded and thus the stretching cannot be arbitrarily extensive without folding. This means the stretched signal will have multiple folds. The ML decoding projects the received signal to a valid codeword with minimum Euclidean distance. Projection onto an erroneous fold results into an anomalous distortion, which introduces a rather notable estimation error. In practical code design, the weak distortion and the anomalous distortion are two competing aspects — lengthening the codeword curve will relieve the weak distortion but will inevitably introduce more folds and a narrower space

between folds, and hence a higher chance for anomalous distortion; likewise, shortening the codeword curve will reduce the chance for anomalous distortion, but increase the weak distortion. The key is to strike a best balance between these competing factors.

Specifically the weak error can be accurately characterized by the CRLB and the anomalous error can be roughly indicated by the BER.

The CRLB for x_0 and y_0 of the mirrored baker system is given in the following, which is also CRLB of single-input baker's code

$$\begin{aligned} CRLB_{x_0}^{mirror} &= -E_{x_0}^{-1}\left\{\left(\frac{\partial^2}{\partial x_0^2}\log p(\mathbf{r}_{1,x},\mathbf{r}_{1,y},\mathbf{r}_{2,x},\mathbf{r}_{2,y}|x_0,y_0)\right)^2\right\} \\ &= \frac{\sigma^2}{\sum_{k=0}^{N-1}a_{1,k}^2+\sum_{k=0}^{N-1}a_{2,k}^2} = \frac{\sigma^2}{\sum_{k=0}^{N-1}2^{2k}+\sum_{k=0}^{N-1}2^{-2k}} \\ &= \frac{3\sigma^2}{4^N-4^{1-N}+3} = CRLB_{y_0}^{mirror} = CRLB_{y_0}^{1-d}. \end{aligned} \tag{3.48}$$

For comparison, CRLB for the tent map code coderate $1/(2N)$ is given as

$$CRLB_{x_0}^{tent} = \frac{3\sigma^2}{4^{2N}-1}. \tag{3.49}$$

It is not hard to verify the fact that

$$CRLB_{x_0}^{tent} < CRLB_{x_0}^{mirror} = CRLB_{x_0}^{1-d}, \forall N \in \mathbb{N}^+. \tag{3.50}$$

This means under equal bandwidth expansion(or coderate), tent map system will always have a lower weak distortion.

For tent map and baker map coding systems, itinerary errors cause anomalous distortion. To compare the anomalous distortion of different analog coding systems, we

examine the bit error rate(BER) performance of the itinerary bits for each code. We test the tent map code, mirrored baker's code and single-input code with $N = 5$, each of which has itinerary length of 4. The BER of each itinerary bit for different systems are illustrated in sub-figures of Fig.3.8. It should be noted that in Fig.3.8, the tent map code has code rate of $1/5$ while mirrored baker's and single-input baker's system has code rate of $1/10$. The BER performance for the first 4 itinerary bits of rate-$1/10$ tent map system are even worse than those for the rate-$1/5$ tent map code.

From the figures in Fig.3.8, the mirrored baker's map code and single-input baker's map code have obvious advantage in the itinerary BER performance. The mirrored structure exhibits equal protection for different itinerary bits and the BER decay with steeper slope than those of tent map code. Comparatively the single-input baker's system presents an unequal protection of different itinerary bits. The BER for itinerary bits with smaller indices decays much faster than those with larger indices. Since errors in itinerary bit with smaller index cause more serious distortion, the single-input baker system performs a clever unequal protection to itinerary bits adaptive to their significance. This also explains the single-input baker's map code's advantageous performance over the mirrored baker's map code in the medium SNR range.

From the above comparison, it can be seen that although the improved baker's analog codes have larger weak distortion than the tent map code, their anomalous distortion has been effectively suppressed. The modified baker's map codes achieve a better balance between the protection against two kinds of distortion and consequently outperform the tent map code in a wide SNR range.

Next we compare baker's map code with optimum performance theoretically attainable(OPTA) and existing analog coding schemes in literature [25, 26, 76]. OPTA can be

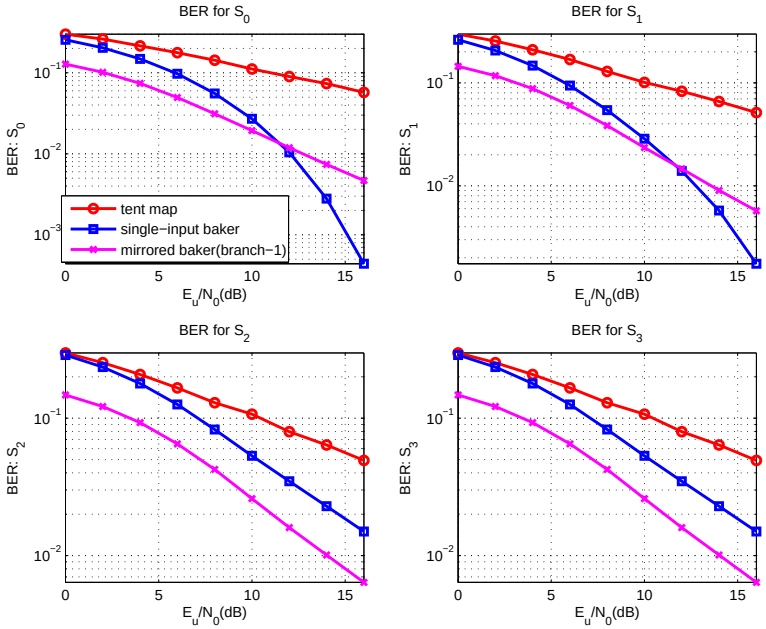

Figure 3.8: BER of itinerary bits for tent map code, mirrored baker's map code and single-input baker's map code($N = 5$)

obtained by equating the rate distortion function with the channel capacity. From [35], we know that the rate distortion function depends on the source distribution and usually does not have a closed-form expression. One of the few exceptions is the Gaussian source, whose distortion function can be obtained analytically (Theorem 13.3.2 in [35]). In the Gaussian case, OPTA can be obtained in a closed form and this is part of the reasons why the existing literature tends to choose Gaussian sources as the case of study, like [24–26, 76] do. However, Gaussian sources can not be fed directly to the family of baker's map encoders, whose inputs are required to be bounded ($[-1, +1]$). Nevertheless to make our proposal comparable with OPTA and other previous works, we perform the comparison in an approximated manner by using truncated Gaussian source. The source signal is first generated from the Gaussian distribution $\mathcal{N}(m, \sigma^2) = \mathcal{N}(0, (1/3)^2)$. We then truncate it using a limiting range of $3\sigma = 1$, such that 99.7% of the probability

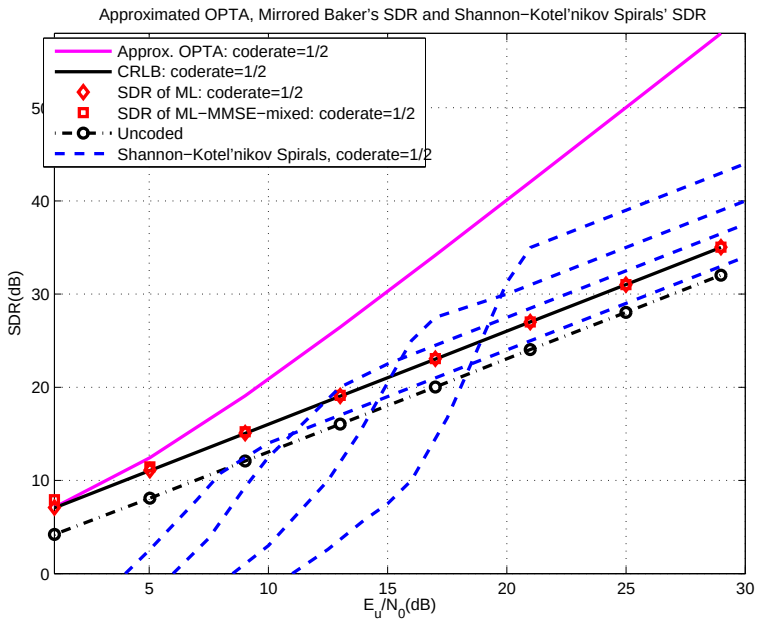

Figure 3.9: Approximated OPTA, SDR of Mirrored Baker's Map Code and Shannon-Kotel'nikov Spirals with Different Parameters

mass falls in the region of $[-1, +1]$. The signal value is set as $+1$ if it exceeds $+1$, and -1 if it drops below -1. We performed mirrored maker's coding on this truncated Gaussian source, and the results are shown in Fig.3.9. It should be noted that in the figure, the OPTA bound is calculated with the true Gaussian source (the only source that is analytically tractable). Since the simulated coding schemes use a truncated Gaussian source, we therefore see a small discrepancy, the baker's code actually appears to slightly outperform the OPTA at the low SNR region. At the same time, we also plot the series of Shannon-Kotel'nikov spirals with parameters optimized for different channel SNR(figure 9 in [25]). It should be noted that the MSE performance of mirrored baker's code and Shannon-Kotel'nikov spirals in Fig.3.9 are obtained by ML method, which can be improved by MMSE method according to [76] and our previous discussion.

The advantage of the parameterized Shannon-Kotel'nikov spiral curve approach is

that by optimizing the parameters with respect to the source distribution and the channel condition, the performance of the code can be made within some 5 dB from the OPTA [25]. The cost, however, is that one must know the exact source distribution and the accurate SNR information. As shown in Fig.3.9, each curve represents a Shannon-Kotel'nikov spiral with its parameter optimized towards one specific channel SNR. Every time the channel condition changes (i.e. a different SNR), the parameter(s) must be adapted, or the code will suffer from a quick performance deterioration due to channel-mismatch.

The proposed baker's analog codes do not require the knowledge of the source distribution nor the channel SNR in order to perform encoding and ML decoding. Instead of designing a sequence of codes, one optimized for each channel SNR in [25], in our approach, a single code is used for a wide range of SNR range. Fig.3.9 reflects that our proposal's SDR(in dB) has identical slope for high channel SNR, or diversity, as those of optimized Shannon-Kotel'nikov spirals. The improved baker's analog codes universally outperforms the Shannon-Kotel'nikov spirals optimized for low channel SNR and has obvious advantage in low SNR range for all Shannon-Kotel'nikov spirals. Additionally, the ML decoding algorithm of our proposed chaotic analog codes has simple closed-form expression, which is absent for spiral codes.

Last we compare the proposed analog encoding system with the conventional digital encoding systems for analog signals transmission. In our experiment, the source signals are uniformly distributed between the range $[-1, +1]$. For digital systems, uniform quantization and turbo codes with recursive systematic convolutional code $(1, \frac{1+D+D^2+D^3}{1+D+D^3})$ are used. The BCJR(log-MAP) algorithm with 8 decoding iterations is performed for decoding the turbo code. Uniform puncturing is utilized to appropriately adjust the

coderate when applicable. Due to the different significance of bits obtained by quantization, equal error protection(EEP) and unequal error protection(UEP) are considered. The details of the tested systems are given as follows:

1. Analog: $(6, 2)$ analog code is used by utilizing the mirrored baker's code with $N = 2$ and puncturing the system signals (y_0, x_0) for the second branch. Assuming that codewords are transmitted using in-phase and quadrature forms(which can be regarded as ∞-QAM modulation), the system has bandwidth expansion of $3/2$.

2. Digital-EEP: 8-bit quantization, $(3072, 2048, 2/3)$ turbo code and 256-QAM are used. System bandwidth expansion is $3/2$.

3. Digital-UEP1: 8-bit quantization is performed. The four least significant bits(LSB) are left uncoded. The four most significant bits(MSB) are encoded by $(4096, 2048, 1/2)$ turbo code. Both the coded and uncoded bits are 256-QAM modulated. System bandwidth is $3/2$.

4. Digital-UEP2: 8-bit quantization is performed. The two LBS are uncoded. The six MSB are encoded by $(3410, 2046, 3/5)$ turbo code. All bits are 256-QAM modulated. System bandwidth is $3/2$.

5. Digital-UEP3: 8-bit quantization is performed. The four LSB are uncoded. The four MSB are encoded by $(2560, 2048, 4/5)$ turbo code. The coded and uncoded bits go through 64-QAM modulation. System bandwidth is $3/2$.

The performance of the proposed analog and four digital systems are plotted in Fig.3.10. The proposed analog code exhibits an obvious advantage to the digital competitors over a wide range when SNR has low and medium values. The digital systems

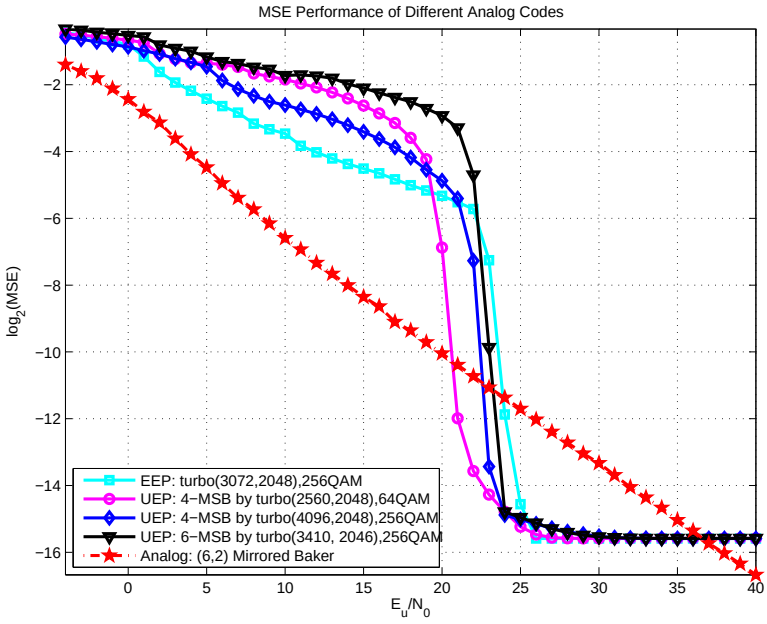

Figure 3.10: Analog signals transmission—analog system vs digital system

enter their waterfall region at rather high SNR and exhibits error floor, which is result of the quantization noise. In fact, due to the bandwidth limitation, quantization noise will always exist for the digital transmission schemes, and eventually form an error floor that limits the overall system perform even as the SNR increases to infinity. In Fig.3.10, the digital coding schemes outperform the analog scheme in a narrow E_u/N_o range, which is due to the fact that the digital error correction codes' performance boosts drastically in a very narrow SNR range (the so-called water-fall region). The digital codes' resilience to noise, although which is powerful, is finally suppressed by the quantization noise. Comparatively, analog coding schemes have a very graceful performance evolution and its distortion can be made arbitrarily small if the channel is sufficiently good.

3.6 Conclusion

This chapter introduces a family of pure analog chaotic dynamic encoding schemes based on baker's map function. We first discuss the coding scheme using original baker's map function, including its encoding and decoding schemes. Mean square error analysis indicates that the intrinsic unbalanced protection of its input results in a unsatisfying performance. Based on that two improvement encoding schemes are proposed—mirrored baker's and single-input baker's system. These two schemes provide sufficient protection to all encoded analog sources. The various decoding methods for the original baker's coding system are extended to the modified systems. Compared to the classical tent map analog code, the improved baker's map encoding schemes achieve a better balance between the anomalous and weak distortion and have advantageous performance in a wide practical SNR range. Moreover, our improved encoding schemes also exhibit competing or even better performance than the classical analog joint source-channel coding scheme, especially in low SNR range, while maintain much lower complexity in decoding procedure. We also compare the analog and conventional digital systems using turbo code to transmit analog source signals. The digital systems suffer from the granularity noise due to quantization, large decoding latency and threshold effect. Comparatively, the analog coding scheme has a graceful performance degradation and outperforms over a wide SNR region.

3.7 Appendix

In this appendix, we provide detailed proof of the closed form solution of MMSE decoder for the original baker's map code in (3.19).

Following notations in the section 3.2, we start from equation (3.15), the MMSE estimate of x_0 can be given as

$$\hat{x}_0^{MMSE} = \mathsf{E}\{x_0|\boldsymbol{r}_x, \boldsymbol{r}_y\} = \int_{-1}^{+1} x_0 f(x_0|\boldsymbol{r}_x, \boldsymbol{r}_y)\mathrm{d}x_0 \tag{3.51}$$

$$= \sum_{\boldsymbol{s}} \int_{C_{\boldsymbol{s}}} x_0 f(x_0|\boldsymbol{r}_x, \boldsymbol{r}_y)\mathrm{d}x_0 \tag{3.52}$$

$$= \sum_{\boldsymbol{s}} \int_{C_{\boldsymbol{s}}} x_0 \frac{f(\boldsymbol{r}_x, \boldsymbol{r}_y|x_0)f(x_0)}{f(\boldsymbol{r}_x, \boldsymbol{r}_y)}\mathrm{d}x_0 \tag{3.53}$$

$$= \frac{1}{2f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \sum_{\boldsymbol{s}} \int_{C_{\boldsymbol{s}}} x_0 \int_{-1}^{+1} f(\boldsymbol{r}_x, \boldsymbol{r}_y|x_0, y_0)f(y_0|x_0)\mathrm{d}y_0\mathrm{d}x_0 \tag{3.54}$$

$$= \frac{1}{4f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \sum_{\boldsymbol{s}} \int_{C_{\boldsymbol{s}}} x_0 \int_{-1}^{+1} f(\boldsymbol{r}_x, \boldsymbol{r}_y|x_0, y_0)\mathrm{d}y_0\mathrm{d}x_0 \tag{3.55}$$

$$= \frac{1}{4f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \sum_{\boldsymbol{s}} \int_{C_{\boldsymbol{s}}} x_0 \int_{-1}^{+1} \Big[\frac{1}{\sqrt{2\pi}\sigma}\Big]^{2N} \exp\Bigg\{ -\frac{1}{2\sigma^2} \sum_{k=0}^{N-1} \Big\{ \big[r_{x,k} - (a_{k,\boldsymbol{s}_n}x_0 + b_{k,\boldsymbol{s}_n})\big]^2 + \big[r_{y,k} - (c_{k,\boldsymbol{s}_n}y_0 + d_{k,\boldsymbol{s}_n})\big]^2 \Big\} \Bigg\} \mathrm{d}y_0\mathrm{d}x_0. \tag{3.56}$$

In the above equations, we utilize the fact that x_0 and y_0 are independently uniformly distributed over the range $[-1, +1]$. To proceed with the above derivation, we introduce some intermediate parameters as follows:

$$\begin{aligned} A_1 &= \|\boldsymbol{a}_{\boldsymbol{s}}\|^2; \quad B_1 = \boldsymbol{a}_{\boldsymbol{s}}^T(\boldsymbol{b}_{\boldsymbol{s}} - \boldsymbol{r}_x); \quad C_1 = \|\boldsymbol{b}_{\boldsymbol{s}} - \boldsymbol{r}_x\|^2; \\ A_2 &= \|\boldsymbol{c}_{\boldsymbol{s}}\|^2; \quad B_2 = \boldsymbol{c}_{\boldsymbol{s}}^T(\boldsymbol{d}_{\boldsymbol{s}} - \boldsymbol{r}_y); \quad C_2 = \|\boldsymbol{d}_{\boldsymbol{s}} - \boldsymbol{r}_y\|^2; \end{aligned} \tag{3.57}$$

Thus the calculation in (3.51) can be further written as

$$\hat{x}_0^{MMSE} = \frac{\left(2\pi\sigma^2\right)^{-N}}{4f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \sum_{\boldsymbol{s}} \Bigg\{ \underbrace{\int_{C_{\boldsymbol{s}n}} x_0 \exp\Big\{ -\frac{1}{2\sigma^2}\big[A_1 x_0^2 + 2B_1 x_0 + C_1\big]\Big\} \mathrm{d}x_0}_{I_1(\boldsymbol{s})} \cdot \underbrace{\int_{-1}^{+1} \exp\Big\{ -\frac{1}{2\sigma^2}\big[A_2 y_0^2 + 2B_2 y_0 + C_2\big]\Big\} \mathrm{d}y_0}_{I_2(\boldsymbol{s})} \Bigg\}. \tag{3.58}$$

Similarly, the MMSE estimator of $\hat{y}_0^{MMSE}$ can be also obtained starting from (3.15) and is determined as

$$\hat{y}_0^{MMSE} = E\{y_0 | \boldsymbol{r}_x, \boldsymbol{r}_y\} = \int_{-1}^{+1} y_0 f(y_0 | \boldsymbol{r}_x, \boldsymbol{r}_y) \mathrm{d}y_0 \tag{3.59}$$

$$= \int_{-1}^{+1} y_0 \int_{-1}^{+1} \frac{f(\boldsymbol{r}_x, \boldsymbol{r}_y | x_0, y_0) f(y_0) f(x_0 | y_0)}{f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \mathrm{d}x_0 \mathrm{d}y_0 \tag{3.60}$$

$$= \frac{1}{4f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \int_{-1}^{+1} y_0 \sum_{n=0}^{2^{N-1}-1} \int_{C_{\boldsymbol{s}n}} f(\boldsymbol{r}_x, \boldsymbol{r}_y | x_0, y_0) \mathrm{d}x_0 \mathrm{d}y_0 \tag{3.61}$$

$$= \frac{\left(2\pi\sigma^2\right)^{-N}}{4f(\boldsymbol{r}_x, \boldsymbol{r}_y)} \sum_{\boldsymbol{s}} \Bigg\{ \underbrace{\int_{-1}^{+1} y_0 \exp\Big\{ -\frac{1}{2\sigma^2}\big[A_2 y_0^2 + 2B_2 y_0 + C_2\big]\Big\} \mathrm{d}y_0}_{I_3(\boldsymbol{s})} \cdot \underbrace{\int_{C_{\boldsymbol{s}n}} \exp\Big\{ -\frac{1}{2\sigma^2}\big[A_1 x_0^2 + 2B_1 x_0 + C_1\big]\Big\} \mathrm{d}x_0}_{I_4(\boldsymbol{s})} \Bigg\}. \tag{3.62}$$

Observing equations (3.58) and (3.62), the term $f(\boldsymbol{r}_x, \boldsymbol{r}_y)$ still needs to be deter-

mined, which can be calculated as

$$f(\boldsymbol{r}_x, \boldsymbol{r}_y) = \int_{-1}^{+1}\int_{-1}^{+1} f(\boldsymbol{r}_x, \boldsymbol{r}_y|x_0, y_0) f(x_0) f(y_0) \mathrm{d}x_0 \mathrm{d}y_0 \tag{3.63}$$

$$= \frac{\left(2\pi\sigma^2\right)^{-N}}{4} \sum_{\boldsymbol{s}} \Big(I_2(\boldsymbol{s}) I_4(\boldsymbol{s}) \Big) \tag{3.64}$$

where $I_2(\boldsymbol{s})$ and $I_4(\boldsymbol{s})$ are defined in (3.58) and (3.62) respectively. Here we further introduce the following notations:

$$\begin{aligned}
E_1 &= \exp\left\{\frac{B_1^2 - A_1C_1}{2\sigma^2 A_1}\right\};\ \ D_1 = Q\bigg(\frac{\sqrt{A_1}}{\sigma} e_{l,\boldsymbol{s}} + \frac{B_1}{\sigma\sqrt{A_1}}\bigg) - Q\bigg(\frac{\sqrt{A_1}}{\sigma} e_{u,\boldsymbol{s}} + \frac{B_1}{\sigma\sqrt{A_1}}\bigg);\\
E_2 &= \exp\left\{\frac{B_2^2 - A_2C_2}{2\sigma^2 A_2}\right\};\ \ D_2 = Q\bigg(-\frac{\sqrt{A_2}}{\sigma} + \frac{B_2}{\sigma\sqrt{A_2}}\bigg) - Q\bigg(\frac{\sqrt{A_2}}{\sigma} + \frac{B_2}{\sigma\sqrt{A_2}}\bigg);\\
J_1 &= \exp\bigg\{-\frac{1}{2\sigma^2}A_1\Big(e_{l,\boldsymbol{s}} + \frac{B_1}{A_1}\Big)^2\bigg\} - \exp\bigg\{-\frac{1}{2\sigma^2}A_1\Big(e_{u,\boldsymbol{s}} + \frac{B_1}{A_1}\Big)^2\bigg\};\\
J_2 &= \exp\bigg\{-\frac{1}{2\sigma^2}A_2\Big(1 + \frac{B_2}{A_2}\Big)^2\bigg\} - \exp\bigg\{-\frac{1}{2\sigma^2}A_2\Big(-1 + \frac{B_2}{A_2}\Big)^2\bigg\},
\end{aligned} \tag{3.65}$$

where the function $Q(\cdot)$ is the well known Gaussian-Q function which is defined as

$$Q(x) \triangleq \int_x^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2}} \mathrm{d}t. \tag{3.66}$$

After some manipulations, the integrals $I_1(\boldsymbol{s})$, $I_2(\boldsymbol{s})$, $I_3(\boldsymbol{s})$ and $I_4(\boldsymbol{s})$ defined previously can be given by use of the notations in (3.17) as

$$\begin{aligned}
I_1(\boldsymbol{s}) &= E_1\bigg(\frac{\sigma^2}{A_1}J_1 - \frac{\sqrt{2\pi}B_1\sigma}{A_1^{3/2}}D_1\bigg);\quad I_2(\boldsymbol{s}) = \sqrt{\frac{2\pi}{A_2}}\sigma E_2 D_2;\\
I_3(\boldsymbol{s}) &= E_2\bigg(\frac{\sigma^2}{A_2}J_2 - \frac{\sqrt{2\pi}B_2\sigma}{A_2^{3/2}}D_2\bigg);\quad I_4(\boldsymbol{s}) = \sqrt{\frac{2\pi}{A_1}}\sigma E_1 D_1;
\end{aligned} \tag{3.67}$$

Thus by substituting the equations (3.57), (3.65) and (3.67) into (3.58) and (3.62), we

can finally obtain the MMSE estimator of x_0 and y_0 as the following

$$\hat{x}_0^{MMSE} = \frac{\sum_{\boldsymbol{s}} \sqrt{\frac{2\pi}{A_2}} E_1 E_2 D_2 \left(\frac{\sigma}{A_1} J_1 - \frac{\sqrt{2\pi} B_1}{A_1^{3/2}} D_1 \right)}{\sum_{\boldsymbol{s}} \frac{2\pi}{\sqrt{A_1 A_2}} E_1 E_2 D_1 D_2}, \tag{3.68}$$

$$\hat{y}_0^{MMSE} = \frac{\sum_{\boldsymbol{s}} \sqrt{\frac{2\pi}{A_1}} E_1 E_2 D_1 \left(\frac{\sigma}{A_2} J_2 - \frac{\sqrt{2\pi} B_2}{A_2^{3/2}} D_2 \right)}{\sum_{\boldsymbol{s}} \frac{2\pi}{\sqrt{A_1 A_2}} E_1 E_2 D_1 D_2}. \tag{3.69}$$

The proof has been completed.

Chapter 4

Nonbandwidth-Expansion Precoding for Sensor Wireless Network Transformation

4.1 Introduction

Along with the advancement of the microelectromechanical systems(MEMS) and wireless communication techniques, the wireless sensor network(WSN) has become a attracting technology for its wide spectrum of application [80]. In the past few years, considerable interest has been cast onto wireless sensor networks and great results have been obtained in both theory and practical implementations.

In spite of its promising applications, wireless sensor network confronts multiple challenges at the same time. A wireless sensor network is typically composed by large

number of geographically distributed and wireless connected sensor nodes. Each sensor node is usually battery-bearing, irretrievable and low-price unit, which is equipped with sensing device to harvest information from the environment, capability-limited processor to partially process the incoming raw signals and transceiver to transmit data or get controlled message to/from the fusion center or neighbors. The prevailing wireless sensor networks have two typical structures: ad hoc network and networks with fusion centers. For the former structure the self-organized routing protocol becomes a problem, and for the latter one, low-latency processing algorithms should are required since the latency at the fusion center will be specially concerned. Besides what is mentioned above, bandwidth is also an important considering in many practical scenarios, such as the emerging wireless multimedia sensor networks(WMSN) [36].

Though the design of wireless sensor network involves multi-discipline viewpoints, this chapter focuses on the wireless communication facet of the problem. For sensor network communication, the degradation in performance mainly comes from fading. So fading-resistant techniques should be employed to improve communication reliability. However due to the concern of cost reduction or limit of size, usually it is impractical to equip the sensor with multiple antennas. Thus it is impossible to utilize the abundant well designed techniques exploiting space diversity, such as space time code(STC) in [46]. Channel error correction code may be a possible way to improve performance. However well-performed codes, like turbo and low-density-priority-code(LDPC) that achieve Shannon's limit, have long latency and thus become impractical for the networks with fusion centers. Moreover, channel coding inevitably expands bandwidth so it is not applicable for the high rate sensor networks like the wireless multimedia sensor networks. Thus a suitable transmission scheme tailored for the practical wireless sensor network deserves discussion. Reference [45] proposed a scheme named signal space

diversity(SSD) to utilize diversity without requiring multiple antennas nor introducing bandwidth expansion. Inspired by that, we propose a communication scheme which utilize unitary precoding before transmission. We will discuss its performance, design criterion and detection methods in the following.

This chapter will be organized as the following. Section 4.2 will talk the system model of the precoded sensor transmission system. Section 4.3 will analyze the system's error probability performance with reasonable approximation and consequently obtain conclusions on the precoder design criteria. Section 4.4 will discuss the detecting algorithm and extensive simulation results will be given in Section 4.5.

Notation: Bold lower(upper) case letters denote column vectors(matrices); $(\cdot)^H$ denotes conjugate transpose. $\|\cdot\|$ denotes the Frobenius norm; $\|\mathbf{b}\|_{\mathbf{A}}$ denotes the $\sqrt{\mathbf{b}^H\mathbf{A}\mathbf{b}}$, where $\mathbf{b} \in \mathbb{C}^{K\times 1}$ and $\mathbf{A} \in \mathbb{C}^{K\times K}$ is positive semi-definite; $\mathfrak{Re}(\cdot)$ denotes the real part of a complex number; $\mathbf{I}_K$ denotes identity matrix of K dimension.

4.2 System Model

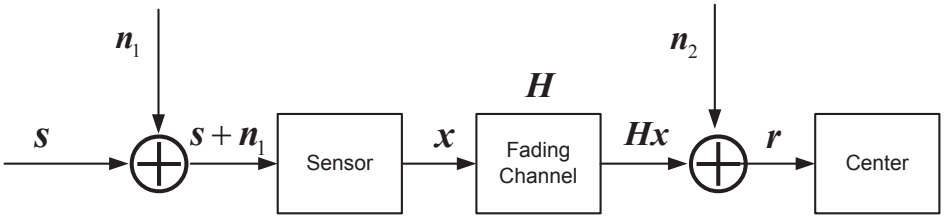

Figure 4.1: The Sensor-Center Wireless Communication System

We focus on the wireless transmission from a sensor node to another sensor/fusion center, as illustrated in Fig.1. Here it is assumed that the sensed signal from the source is binary signal, i.e. $u_i \in \{\pm\sqrt{E_s}\}$, which can serve as the indication of occurrence of

some specific events being monitored. By packing every pair of sensed data $\{u_{2i-1}, u_{2i}\}$ into one complex number, we express incoming data as s_i. Due to the thermal noise of devices or interference from the background environment, the incoming signal may often be contaminated in practice. Thus it is reasonable to model this disturbance at input as an circular symmetric Gaussian noise $n_{1i} \sim \mathcal{CN}(0, \sigma_1^2)$. The obtained data is then transmitted to the destination node through a wireless communication channel and is attenuated by an fading factor h_i. Here we assume that the channel is frequency nonselective and fast fading channel. A typical and easy model of such channel can be i.i.d. Rayleigh fading and thus $h_i \sim \mathcal{CN}(0, 1)$. At the destination, the transmitted signal from sensor is, again, corrupted by circular symmetric Gaussian noise $n_2 \sim \mathcal{CN}(0, \sigma_2^2)$.

By packing every consecutive K pairs of incoming data into a complex vector, the system model stated in the last subsection can be compactly expressed as the following:

$$\mathbf{r} = \mathbf{H}\mathbf{G}^H(\mathbf{s} + \mathbf{n}_1) + \mathbf{n}_2 \tag{4.1}$$

$$= \mathbf{H}\mathbf{G}^H\mathbf{s} + (\mathbf{H}\mathbf{G}^H\mathbf{n}_1 + \mathbf{n}_2) \tag{4.2}$$

$$= \mathbf{H}\mathbf{G}^H\mathbf{s} + \mathbf{n}, \tag{4.3}$$

where $\mathbf{s}, \mathbf{n}_1$ and $\mathbf{n}_2 \in \mathbb{C}^{K\times 1}$ are the original sensed signal, noise at the sensor and noise at the destination respectively. $\mathbf{H} \in \mathbb{C}^{K\times K}$ is the diagonal channel matrix, with its diagonal elements h_i being channel fading coefficients for different time instances. $\mathbf{G} \in \mathbb{C}^{K\times K}$ is the precoder to be designed.

Here to make the following discussion easy and clear, some assumptions are adopted:

A1) For simplicity it is assumed that $-\sqrt{E_s}$ and $+\sqrt{E_s}$ are equally probable.[1] *The*

[1]Actually this may not be true in practice, but still we can adjust positive and negative amplitudes to

signals corresponding to different time instances are independent. This can be compactly expressed as $E[\mathbf{s}] = \mathbf{0}$ *and* $E[\mathbf{s}^H\mathbf{s}] = E_s\mathbf{I}_N$.

A2) $\mathbf{G}$ *is invertible, so that the transmitted signal* $\mathbf{s}$ *can perfectly recovered when the noise is absent.*

A3) The average transmission power is constant, that is

$$E\{\|\mathbf{G}^H\mathbf{s}\|^2\} = E_s tr\Big(\mathbf{G}^H\mathbf{G}\Big) = E_s tr\Big(\mathbf{G}\mathbf{G}^H\Big) = P. \tag{4.4}$$

A4) It is assumed that the channel status information at the receiver(CSIR), i.e. $\mathbf{H}$ *is available at the center.*

4.3 Precoder Design

Based on the above description, the problem has come up that the precoder $\mathbf{G}^H$ should be carefully designed aiming at a good system performance. Since symbol error probability(SEP) is a standard norm for assessing performance, it is natural to investigate the design criterion for $\mathbf{G}$ by pairwise error probability(PEP) analysis. The disturbance at the sensor makes the noise $\mathbf{n}$ at the destination node colored and relevant of the fading $\mathbf{H}$, which case is seldom discussed in the classical point-to-point wireless communication and makes the analysis difficult, as will soon be seen.

make its mean zero.

4.3.1 PEP Analysis

Assume that the optimal detecting, i.e. Maximum Likelihood(ML) detecting, is performed at the destination node. In the following, the PEP $P\big(\mathbf{s}_0 \to \mathbf{s}_1\big)$ is considered, where $\mathbf{s}_0$ is actually transmitted and the decision $\hat{\mathbf{s}}$ is made as $\mathbf{s}_1$ instead of $\mathbf{s}_0$.

First according to (4.3), it is obvious that $\mathbf{n}$ is zero mean complex Gaussian noise. Generally $\mathbf{n}$ is colored noise. Its covariance matrix $\mathbf{\Sigma_n}$ can be calculated as the following:

$$\mathbf{\Sigma_n} = E[\mathbf{n}\mathbf{n}^H] \tag{4.5}$$

$$= E\Big[\Big(\mathbf{H}\mathbf{G}^H\mathbf{n}_1 + \mathbf{n}_2\Big)\Big(\mathbf{n}_1^H\mathbf{G}\mathbf{H}^H + \mathbf{n}_2^H\Big)\Big] \tag{4.6}$$

$$= \sigma_1^2\mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H + \sigma_2^2\mathbf{I}_K \tag{4.7}$$

Since $\mathbf{H}$ is known at the center, the likelihood function $p(\mathbf{r}|\mathbf{s},\mathbf{H})$ can be given as:

$$p(\mathbf{r}|\mathbf{s},\mathbf{H}) = \det\big(\pi\mathbf{\Sigma_n}\big)^{-1}\exp\Big[-(\mathbf{r}-\mathbf{H}\mathbf{G}^H\mathbf{s})^H\mathbf{\Sigma}_\mathbf{n}^{-1}(\mathbf{r}-\mathbf{H}\mathbf{G}^H\mathbf{s})\Big]$$

So the ML detection result $\hat{\mathbf{s}}_{ML}$ should be given as:

$$\hat{\mathbf{s}}_{ML} = \min_{\mathbf{s}}\|\mathbf{r}-\mathbf{H}\mathbf{G}^H\mathbf{s}\|^2_{\mathbf{\Sigma}_\mathbf{n}^{-1}} \tag{4.8}$$

The conditional probability of pairwise error $P(\mathbf{s}_0 \to \mathbf{s}_1|\mathbf{H})$ can be approximately writ-

ten as

$$P(\mathbf{s}_0 \to \mathbf{s}_1|\mathbf{H}) \approx P\Big(\|\mathbf{r}-\mathbf{H}\mathbf{G}^H\mathbf{s}_1\|^2_{\mathbf{\Sigma}_\mathbf{n}^{-1}} < \|\mathbf{r}-\mathbf{H}\mathbf{G}^H\mathbf{s}_0\|^2_{\mathbf{\Sigma}_\mathbf{n}^{-1}}\Big)$$
$$= P\Big(2\mathfrak{Re}\{\mathbf{n}^H\mathbf{\Sigma}_\mathbf{n}^{-1}\mathbf{H}\mathbf{G}^H(\mathbf{s}_0-\mathbf{s}_1)\} < -\|\mathbf{H}\mathbf{G}^H(\mathbf{s}_0-\mathbf{s}_1)\|^2_{\mathbf{\Sigma}_\mathbf{n}^{-1}}\Big)$$

Since the left hand side of the above equation follows the Gaussian distribution $\mathcal{N}\big(0, 2\|\mathbf{H}\mathbf{G}^H(\mathbf{s}_0-\mathbf{s}_1)\|^2_{\mathbf{\Sigma}_\mathbf{n}^{-1}}\big)$, it can be further simplified into a compact form

$$P\Big(\mathbf{s}_0 \to \mathbf{s}_1|\mathbf{H}\Big) = Q\Big(\sqrt{\frac{1}{2}\|\mathbf{H}\mathbf{G}^H(\mathbf{s}_0-\mathbf{s}_1)\|^2_{\mathbf{\Sigma}_\mathbf{n}^{-1}}}\Big) \tag{4.9}$$

where $Q(\cdot)$ is the well-known Gaussian-Q function. The PEP $P\big(\mathbf{s}_0 \to \mathbf{s}_1\big)$ should be obtained by averaging the conditional PEP over all possible realization of channel states, that is

$$P\big(\mathbf{s}_0 \to \mathbf{s}_1\big) = E_\mathbf{H}\Big[P(\mathbf{s}_0 \to \mathbf{s}_1|\mathbf{H})\Big]. \tag{4.10}$$

Thus the problem of design of $\mathbf{G}$ under PEP criterion has become an optimization problem aiming to minimizing the objective in (4.10) with the constraint by (4.4). However noticing that $\mathbf{\Sigma}_\mathbf{n}$ is a function of $\mathbf{H}$ and $\mathbf{G}$ and appears in equation (4.9) in a inverse form, this problem is rather difficult and may not have closed form.

4.3.2 Discussion—High and Low SNR Cases

Since it seems not easy to get some explicit mathematical form of the design of $\mathbf{G}$, we turn to PEP performance corresponding to some special scenarios by introducing reasonable approximation. In the following, we discuss two unilateral cases: low and

high SNR of the channel. Observing that it is rare case that h_i is zero, it is assumed that $\mathbf{H}$ is invertible.

CASE I: High SNR

In this case, σ_2^2 is small enough compared to σ_1^2, then the noise covariance in equation (4.7) can be approximately written as

$$\mathbf{\Sigma_n} = \sigma_1^2 \mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H + \sigma_2^2\mathbf{I}_N \tag{4.11}$$

$$\approx \sigma_1^2 \mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H. \tag{4.12}$$

Substituting the above into the PEP calculation and by defining $\mathbf{\Delta} = \mathbf{s}_0 - \mathbf{s}_1$, we have

$$\|\mathbf{H}\mathbf{G}^H(\mathbf{s}_0 - \mathbf{s}_1)\|^2_{\mathbf{\Sigma_n}^{-1}} = tr\Big\{\mathbf{\Delta}^H\mathbf{G}\mathbf{H}^H\mathbf{\Sigma}_\mathbf{n}^{-1}\mathbf{H}\mathbf{G}^H\mathbf{\Delta}\Big\} \tag{4.13}$$

$$\approx \frac{1}{\sigma_1^2} tr\Big\{\mathbf{\Delta}^H\mathbf{G}\mathbf{H}^H\mathbf{H}^{-H}\big(\mathbf{G}^H\mathbf{G}\big)^{-1}\mathbf{H}^{-1}\mathbf{H}\mathbf{G}^H\mathbf{\Delta}\Big\} \tag{4.14}$$

$$= \frac{1}{\sigma_1^2}\|\mathbf{s}_0 - \mathbf{s}_1\|^2 \tag{4.15}$$

which is independent of $\mathbf{G}$ and $\mathbf{H}$.

Substitute the equation (4.15) into (4.9) and then into (4.10), we can obtain

$$P\big(\mathbf{s}_0 \to \mathbf{s}_1\big) \approx Q\Big(\sqrt{\frac{1}{2\sigma_1^2}\|\mathbf{s}_0 - \mathbf{s}_1\|^2}\Big). \tag{4.16}$$

The above result suggests that with the increasing of SNR, the effect of precoder $\mathbf{G}$ will vanish and the system performance will be bottlenecked by the intrinsic noise $\mathbf{n}_1$ within the incoming data at the sensor, which will result in an error floor in the BER performance.

CASE II: Low SNR

In this case, σ_2^2 is at least modestly large compared to σ_1^2. By defining $\mathbf{g}_i^T$ as the i-th row of $\mathbf{G}^H$ and $\mathbf{\Delta} = \mathbf{s}_0 - \mathbf{s}_1$, we focus on the argument of the Q function in (4.9)

$$\|\mathbf{H}\mathbf{G}^H\mathbf{\Delta}\|^2_{\mathbf{\Sigma_n^{-1}}} = tr\left(\mathbf{\Sigma_n^{-1}}\mathbf{H}\mathbf{G}^H\mathbf{\Delta}\mathbf{\Delta}^H\mathbf{G}\mathbf{H}^H\right) \tag{4.17}$$

Noticing that $\mathbf{H}$ is a diagonal matrix,

$$\|\mathbf{H}\mathbf{G}^H\mathbf{\Delta}\|^2_{\mathbf{\Sigma_n^{-1}}} = tr\left\{\mathbf{\Sigma_n^{-1}}\begin{bmatrix} h_1\mathbf{g}_1^T\mathbf{\Delta} \\ h_2\mathbf{g}_2^T\mathbf{\Delta} \\ \vdots \\ h_K\mathbf{g}_K^T\mathbf{\Delta} \end{bmatrix}\left[h_1^*(\mathbf{g}_1^T\mathbf{\Delta})^*, \cdots, h_K^*(\mathbf{g}_K^T\mathbf{\Delta})^*\right]\right\}$$

According to matrix theory, if $\mathbf{A}$ and $\mathbf{B}$ are both positive semi-definite matrices with dimension N, then

$$\sum_{i=1}^{N}\lambda_i(\mathbf{A})\lambda_{N-i+1}(\mathbf{B}) \le tr\{\mathbf{AB}\} \le \sum_{i=1}^{N}\lambda_i(\mathbf{A})\lambda_i(\mathbf{B}) \tag{4.18}$$

where $\lambda_1(\mathbf{A}) \ge \lambda_2(\mathbf{A}) \ge \cdots \ge \lambda_N(\mathbf{A})$ are ordered eigenvalues of matrix $\mathbf{A}$ and similar definitions are given to $\lambda_j(\mathbf{B})$. Here by choosing $\mathbf{A}$ as $\mathbf{\Sigma_n^{-1}}$ and $\mathbf{B}$ as

$$\begin{bmatrix} h_1\mathbf{g}_1^T\mathbf{\Delta} \\ h_2\mathbf{g}_2^T\mathbf{\Delta} \\ \vdots \\ h_N\mathbf{g}_2^T\mathbf{\Delta} \end{bmatrix}\left[h_1^*(\mathbf{g}_1^T\mathbf{\Delta})^*, h_2^*(\mathbf{g}_2^T\mathbf{\Delta})^*, \cdots, h_N^*(\mathbf{g}_N^T\mathbf{\Delta})^*\right]. \tag{4.19}$$

Noting that $\mathbf{B}$ is rank one matrix, which means $\lambda_1(\mathbf{B}) = \sum_{i=1}^{K}|h_i|^2|\mathbf{g}_i^T\mathbf{\Delta}|^2$ and $\lambda_2(\mathbf{B}) =$

$\lambda_3(\mathbf{B}) = \cdots = \lambda_K(\mathbf{B}) = 0$, we readily have

$$\lambda_1(\mathbf{B})\lambda_K(\mathbf{\Sigma_n^{-1}}) \leq \|\mathbf{H}\mathbf{G}^H\mathbf{\Delta}\|^2_{\mathbf{\Sigma_n^{-1}}} \leq \lambda_1(\mathbf{B})\lambda_1(\mathbf{\Sigma_n^{-1}}) \tag{4.20}$$

By equation (4.7) we have $\mathbf{\Sigma_n} = \sigma_1^2\mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H + \sigma_2^2\mathbf{I}_K \succcurlyeq \sigma_2^2\mathbf{I}_K$[2], due to the fact that $\mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H$ is positive semi-definite. According to matrix theory, we have

$$\lambda_i(\mathbf{\Sigma_n}) \geq \lambda_i(\sigma_2^2\mathbf{I}) = \sigma_2^2, \quad i \in \{1, 2, \cdots, K\}. \tag{4.21}$$

Thus the eigenvalues of $\mathbf{\Sigma_n}$ can be bounded as:

$$0 \leq \lambda_K(\mathbf{\Sigma_n^{-1}}) \leq \lambda_1(\mathbf{\Sigma_n^{-1}}) \leq \sigma_2^{-2} \tag{4.22}$$

If σ_2^2 is sufficiently large, σ_2^{-2} is small enough and thus by the above equation both $\lambda_K(\mathbf{\Sigma_n^{-1}})$ and $\lambda_1(\mathbf{\Sigma_n^{-1}})$ are bounded in the small range $(0, \sigma_2^{-2}]$. Thus by (4.20) it is reasonable approximation of $\|\mathbf{H}\mathbf{G}^H(\mathbf{\Delta})\|^2_{\mathbf{\Sigma_n^{-1}}}$ by taking the midpoint of the interval $(0, \sigma_2^{-2})$, that is

$$\|\mathbf{H}\mathbf{G}^H(\mathbf{\Delta})\|^2_{\mathbf{\Sigma_n^{-1}}} \approx \frac{\sigma_2^{-2}}{2}\sum_{i=1}^{K}|h_i|^2|\mathbf{g}_i^T\mathbf{\Delta}|^2 \tag{4.23}$$

Thus the conditional PEP is approximated as

$$P\Big(\mathbf{s}_0 \to \mathbf{s}_1|\mathbf{H}\Big) \approx Q\Big(\sqrt{\frac{\sigma_2^{-2}}{4}\big(\sum_{i=1}^{K}|h_i|^2|\mathbf{g}_i^T\mathbf{\Delta}|^2\big)}\Big). \tag{4.24}$$

When σ_2^2 is modestly large, though obtaining a closed form of optimal solution of the

[2] $\mathbf{C} \succcurlyeq \mathbf{D}$ means $\mathbf{C} - \mathbf{D}$ is positive semi-definite.

original problem is hard, we can at least obtain a lower bound of the conditional PEP as

$$P\Big(\mathbf{s}_0 \to \mathbf{s}_1|\mathbf{H}\Big) \geq Q\Big(\sqrt{\frac{\sigma_2^{-2}}{2}\big(\sum_{i=1}^{K}|h_i|^2|\mathbf{g}_i^T\mathbf{\Delta}|^2\big)}\Big) \tag{4.25}$$

where the fact that $Q(\cdot)$ is monotonically decreasing is utilized. Averaging the lower bound in (4.25) over all channel realization

$$P\Big(\mathbf{s}_0 \to \mathbf{s}_1\Big) \geq E_{\mathbf{H}}\left[Q\Big(\sqrt{\frac{\sigma_2^{-2}}{2}\big(\sum_{i=1}^{K}|h_i|^2|\mathbf{g}_i^T\mathbf{\Delta}|^2\big)}\Big)\right] \tag{4.26}$$

$$= \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{i=1}^{K}\frac{1}{1+\frac{\sigma_2^{-2}}{4}|\mathbf{g}_i^T\mathbf{\Delta}|^2\sin^2\theta}\mathrm{d}\theta \tag{4.27}$$

where the identity from (4.26) to (4.27) follows the Craig's formula in [43] and general MGF methods in [44], with details omitted here.

Up to here, under the assumption that channel has at most modestly large SNR, the problem of minimizing PEP can be transformed to a problem minimizing averaging right hand of (4.24) and (4.25) over all realization of $\mathbf{H}$, which has a closed form of (4.27). Actually this problem coincides with the linear constellation precoding design problem presented in [47]. Thus the concluded design criteria in [47] can be borrowed here which are detailed as:

C1) Diversity Gain: $\mathbf{G}^H\mathbf{s}_0$ *should differ with* $\mathbf{G}^H\mathbf{s}_1$ *in every coordinates for any pair* $\{\mathbf{s}_0, \mathbf{s}_1\}$ *that* $\mathbf{s}_0 \neq \mathbf{s}_1$.

C2) Coding Gain: $\mathbf{G}$ *should maximize* $\min_{\mathbf{s}_0\neq\mathbf{s}_1}\prod_{i=1}^{K}\|\mathbf{g}_i^T(\mathbf{s}_0 \neq \mathbf{s}_1)\|$ *over all matrices satisfying C1).*

Summarizing the above cases *I and II*, since PEP performance rarely depends on $\mathbf{G}$ when SNR is high and optimal $\mathbf{G}$ approximately follow on criteria *C1)* and *C2)* when SNR is modestly low. So generally the well designed $\mathbf{G}$ should follow the criteria *C1)* and *C2)* above.

However noticing that finding out the optimal matrices satisfying above criteria needs to solve an optimization problem which is non-convex, thus a closed form solution is still impossible. Since criterion *C1)* guarantees the diversity gain, which is the main contribution to performance improvement, the criteria can be loosed to just *C1)*. As proved in the appendix of [47], there always exists unitary matrices which can guarantee *C1)*. Besides, unitary precoder $\mathbf{G}$ has other merits: i) unitary matrices can achieve the lower bound of mean squared error(MSE) over AWGN channel, as proved in [38], i.e. unitary encoder $\mathbf{G}$ is the optimal linear encoder over AWGN channel; ii) when $\mathbf{G}$ is unitary, the covariance matrix $\mathbf{\Sigma_n}$ in (4.7) reduces to diagonal matrices thus the complexity of solving linear system will reduce from $O(N^3)$ to $O(N)$, which is especially desirable for fusion center, which may have to perform detection of the data from a bunch of sensors. Two kinds of algebraic structure unitary matrix is proposed in [**?**]. However by extensive simulation, it has been found usually the randomly generated unitary $\mathbf{G}$ will have competing or even better BER performance than the class of algebraic constructions, which will be discussed in details in Section 4.5 later. But still these specially designed algebraic structures in [47] enjoy the merits of elegant mathematical representations and easiness for assessing performance, which is desirable in theory and application.

To conclude this section, the optimal precoder $\mathbf{G}$ should follow the the criteria *C1)* and *C2)* above. Though the method to find the exact realization of these criteria still re-

mains an open problem, it is recommended here that randomly generated square unitary matrices can be used which can usually provide decent performance for sensor transmission.

4.4 Detection Scheme at the Destination

4.4.1 Exact and Nearly Exact Detection Methods

Once the designed $\mathbf{G}$ is fixed, the detection scheme at the detection node should be considered. The optimal detection scheme is the ML detection which is presented in (4.8). Since $\mathbf{s}$ is discrete, the genuine ML detection should be obtained by exhaustively searching over all the 4^K candidate codewords. Thus exponentially increasing complexity rapidly makes ML detection prohibitive in practice.

As an alternative of the genuine ML detecting, sphere decoding(SD) [40] [41] is a promising suboptimal detecting algorithm which can achieve near optimal performance. Instead of enumerating all the possible high-dimension codewords, SD algorithm constraints all the candidates within a sphere which is centered around some starting point, which is usually a good guess of the transmitted signal. Though SD has near ML performance, its detecting complexity still remains very high. According to [42] its complexity is approximately $O(K^6)$. Another drawback of SD algorithm is that its detecting delay is not constant, the complexity in the worst case can be significantly higher than average.

Considering in the sensor system, the sensor node/fusion center usually has to process the transmitted signals from multiple sensors, thus sub-optimal algorithms with

lower complexity are desirable.

4.4.2 LMMSE Round Off

One alternative sub-optimal detection scheme with low complexity is linear minimum mean squared error(LMMSE) estimation with round off. Though the transmitted signal is discrete in value, we regard it as continuous and take the mean square error(MSE) as performance criterion. Linear estimation is first performed and then estimate is rounded off to the nearest high dimensional signal point. Defining estimation MSE as

$$\mathcal{E} = tr\{E[(\mathbf{s}-\hat{\mathbf{s}})(\mathbf{s}-\hat{\mathbf{s}})^H]\} \tag{4.28}$$

$$= tr\{E\big[\big(\mathbf{s}-\mathbf{F}(\mathbf{H}\mathbf{G}^H\mathbf{s}+\mathbf{n})\big)\big(\mathbf{s}-\mathbf{F}(\mathbf{H}\mathbf{G}^H\mathbf{s}+\mathbf{n})\big)^H\big]\}, \tag{4.29}$$

and setting differential $\frac{\partial \mathcal{E}}{\partial \mathbf{F}^*} = \mathbf{0}$, optimal LMMSE estimator $\mathbf{F}$ is readily given as

$$\mathbf{F} = E_s\mathbf{G}\mathbf{H}^H\big(E_s\mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H + \mathbf{\Sigma_n}\big)^{-1} \tag{4.30}$$

Specially when $\mathbf{G}$ is constructed as unitary, $\mathbf{F}$ above can be further simplified as

$$\mathbf{F} = \mathbf{G}\mathbf{H}^H\Big((1+\frac{\sigma_1^2}{E_s})\mathbf{H}\mathbf{H}^H + \sigma_2^2\mathbf{I}\Big)^{-1} \tag{4.31}$$

Noticing that the channel $\mathbf{H}$ in our problem is a diagonal matrix, thus the inverse in the above equation just reduces to inverse of K scalars. So the LMMSE estimation here enjoys a very low complexity in our problem. Once the linear optimal estimation is performed, the decision is obtained by rounding off the each coordinate of linear

estimate to the nearest signal point in the constellation.

4.4.3 Partially Nulling and Canceling Method

Although having a very low complexity, LMMSE round off suffers a significant performance degradation. Here we introduce a simple detecting scheme named partially nulling and canceling(PNC) algorithm to perform the detection, which lies between the unsatisfying LMMSE round off and expensive SD algorithm.

When the signal $\mathbf{s}$ has large dimension, we first decode its "strongest" coordinates to reduce the signal dimension. After the signal dimension is modestly decreased, exhaustive search is utilized for the remaining undetected coordinates. Intuitively, the "strongest" coordinates must be the ones with smallest estimation error, which can be presented by the MSE of the LMMSE estimation. Thus the partially nulling and canceling algorithm can be described as follows:
First we perform the linear MMSE detection to the receiver $\mathbf{r}$ as

$$\hat{\mathbf{s}}_{MMSE} = \mathbf{Fr} = E_s \mathbf{G}\mathbf{H}^H \left(E_s \mathbf{H}\mathbf{G}^H \mathbf{G}\mathbf{H}^H + \mathbf{\Sigma_n}\right)^{-1} \mathbf{r}. \tag{4.32}$$

The covariance matrix of the above MMSE estimation can be easily given as

$$\mathbf{P} = E\{(\mathbf{s} - \hat{\mathbf{s}}_{MMSE})(\mathbf{s} - \hat{\mathbf{s}}_{MMSE})^H\} \tag{4.33}$$

$$= \left(E_s^{-1}\mathbf{I} + \mathbf{G}\mathbf{H}^H \mathbf{\Sigma_n^{-1}} \mathbf{H}\mathbf{G}^H\right)^{-1}. \tag{4.34}$$

When $\mathbf{P}$ is calculated, the "strongest" coordinate is the one which has the smallest MMSE estimation error, i.e. the dimension corresponding to the minimal P_{ii}. Then

the MMSE estimate of the "strongest" coordinate is rounded off, detected and removed from the receiver $\mathbf{r}$. Since the signal model in (4.3) can be equivalently written as

$$\mathbf{r} = \mathbf{H}\mathbf{G}^H\mathbf{s} + \mathbf{n} = \sum_{i=1}^{K} \mathbf{H}\underline{\mathbf{g}}_i s_i + \mathbf{n} \tag{4.35}$$

with $\underline{\mathbf{g}}_i$ indicating the i-th column of $\mathbf{G}^H$. Once the index of the coordinate with minimal MSE, say k, is determined, the remaining (K-1) coordinates with $\hat{s}_k$ canceled can be compactly expressed as

$$\mathbf{r}_{K-1} = \mathbf{H}\mathbf{G}_{K-1}^H\mathbf{s}_{K-1} + \mathbf{n} \tag{4.36}$$

where $\mathbf{r}_{K-1} = \mathbf{r} - \mathbf{H}\underline{\mathbf{g}}_k\hat{s}_k$, $\mathbf{G}_{K-1}^H = [\underline{\mathbf{g}}_1, \cdots, \underline{\mathbf{g}}_{k-1}, \underline{\mathbf{g}}_{k+1}, \cdots, \underline{\mathbf{g}}_K]$ and $\mathbf{s}_{K-1} = [s_1, \cdots, s_{k-1}, s_{k+1}, s_K]^T$. The same procedure can be performed to the deflated system in (4.36) to extract the second strongest eliminate. This procedure is repeated until the M "strongest" coordinates out of N-dimension $\mathbf{s}$ have been rounded off and canceled from the receiver $\mathbf{r}$. Then the remaining $N-M$ undetected coordinates of $\mathbf{s}$ are determined by exhaustive search. Usually the reasonable choice of M is a consideration of complexity as well as performance.

In the above description, there seems a confusion of the "MMSE". Since during the iterative nulling and canceling procedure, one coordinate is estimated and rounded off in each iterative, the "MMSE" should mean the minimal mean estimated error of each individual coordinate of $\mathbf{r}$, while not the *sum* of mean squared error of all coordinates, which is used in (4.30) to derive LMMSE. However the fact is under some mild assumption, the linear MMSE estimator corresponding to whole vector coincides with that for every individual coordinates. This can be proved in the following

Theorem 4.4.1. When coordinates of $\mathbf{s}$ are independent and zero mean, the linear MMSE estimator that minimizes the MSE of the whole vector coincides with that minimizes MSE of each seperate coordinate.

Proof: Denote s_i as the i-th coordinate of the signal $\mathbf{s}$. According to the hypothesis

$$\mathbf{R}_s = E\{\mathbf{s}\mathbf{s}^H\} = E_s\mathbf{I}_N. \tag{4.37}$$

Assume that the optimal linear MMSE estimator of s_i is $\mathbf{w}_i$, then the MSE associated with the i-th coordinate can be written as

$$\mathcal{E}_i = E\{(\mathbf{w}_i^H\mathbf{r} - s_i)(\mathbf{w}_i^H\mathbf{r} - s_i)^H\} \tag{4.38}$$

to determine $\mathbf{w}_i$, set $\frac{\partial}{\partial \mathbf{w}_i^*}\mathcal{E}_i = 0$ and use (4.37) we have

$$\mathbf{w}_i = E_s\big(E_s\mathbf{H}\mathbf{G}^H\mathbf{G}\mathbf{H}^H + \mathbf{\Sigma}_\mathbf{n}\big)^{-1}\mathbf{H}\underline{\mathbf{g}}_i \tag{4.39}$$

where $\underline{\mathbf{g}}_i$ is the i-th column of the matrix $\mathbf{G}^H$.

Thus if we pack the estimators $\mathbf{w}_i^H$ for all i in rows, the MMSE estimator for each

separate coordinate can be written as

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_1^H \\ \mathbf{w}_2^H \\ \vdots \\ \mathbf{w}_K^H \end{bmatrix} = E_s \mathbf{G}\mathbf{H}^H \left(E_s \mathbf{H}\mathbf{G}^H \mathbf{G}\mathbf{H}^H + \mathbf{\Sigma_n} \right)^{-1} \tag{4.40}$$

which is equivalent to MMSE estimator in (4.30). Thus the theorem is proved. It should be noted that when $\mathbf{R_s} \neq E_s\mathbf{I}$, $\mathbf{W}$ is not identical with $\mathbf{F}$.

4.5 Numerical Results

In this section, numerical results are provided.

In Fig.4.2, the scenario where the noise at the destination node is relatively small compared to that at the sensor is studied. The SNR at the sensor is fixed at 8dB. The BER performance associated with precoded and non-precoded transmission systems is illustrated. Among the precoded schemes, a group of BER curves corresponding to randomly generated unitary matrices of dimension 4×4 are given. The well designed 4×4 algebraic structured precoder in [**?**] is also tested here [3]. As illustrated in the figure, the algebraic structured precoder has the same performance with that of random ones. With increase of SNR over the wireless channel, the difference between the precoded and unprecoded schemes tends to vanish and an error floor appears, which is suggested by the high SNR case analysis in Section 4.3.

[3] For 4×4 matrix, the construction A and construction B in [**?**] coincide with each other.

In Fig.4.3 and Fig.4.4, the scenario where the wireless channel SNR is relatively low compared to SNR at the sensor is studied. The SNR at sensor is fixed at 20dB. For either figure, a group of randomly generated unitary matrices of dimension 4×4 or 5×5 with genuine ML detection are tested. Correspondingly, the 4×4 and 5×5 algebraic structured precoders in [**?**][4] is also tested as a benchmark. Also a group of 8×8 or 10×10 randomly generated unitary matrices are testes with both LMMSE and half-PNC for comparison. Significant diversity gain in performance of precoded systems can be obviously observed compared with the non-precoded systems. The 5×5 algebraic structured matrix has almost identical with those randomly generated ones of the same dimension. While for 4×4 algebraic matrix, a performance degradation can be observed by comparing to 4×4 random ones. By comparing $4/8$-PNC v.s. 4×4 ML and $5/10$-PNC v.s. 5×5 ML, it can be seen that approximately 1dB gain is obtained for PNC over the corresponding half dimension ML detection. Moreover it should be noted that $4/8$-PNC has only half of the complexity for 4×4 ML scheme(the complexity of nulling and canceling is very low and thus ignored), as is similar to $5/10$-PNC v.s. 5×5 ML case. This can be explained that the PNC partially exploits the significant diversity gain out of high dimension signals and make a good trade off between the complexity and performance.

4.6 Conclusion

In this chapter, we have focused on the wireless transmission scheme using rate-1 precoder in between nodes in sensor network. Performance is studied from the PEP viewpoint and precoder design criteria have been obtained. When SNR is good enough,

[4]The 5×5 precoder uses the construction A in [**?**].

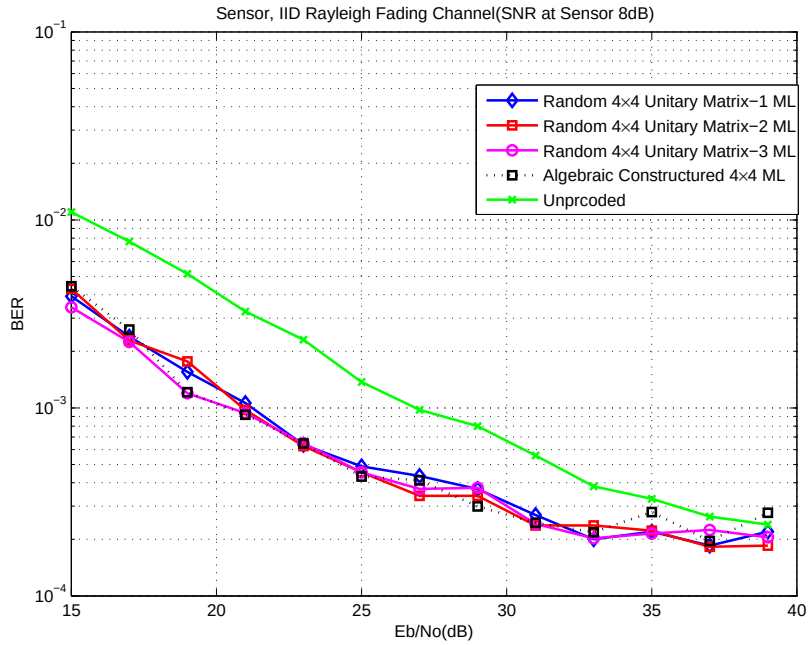

Figure 4.2: SNR at Sensor: 8dB, 4×4 unitary matrix

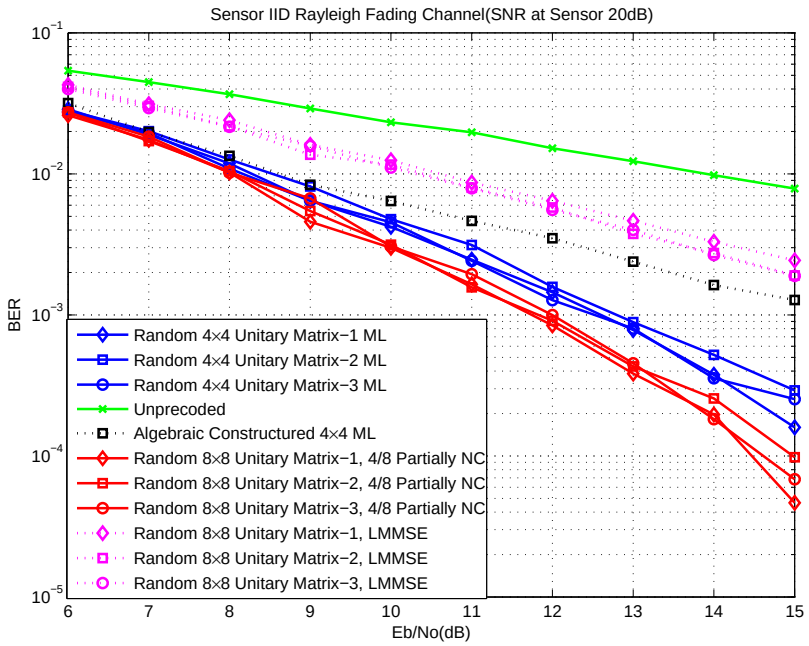

Figure 4.3: SNR at Sensor: 20dB, 4×4 ML and $4/8$ Partially Canceling and Nulling

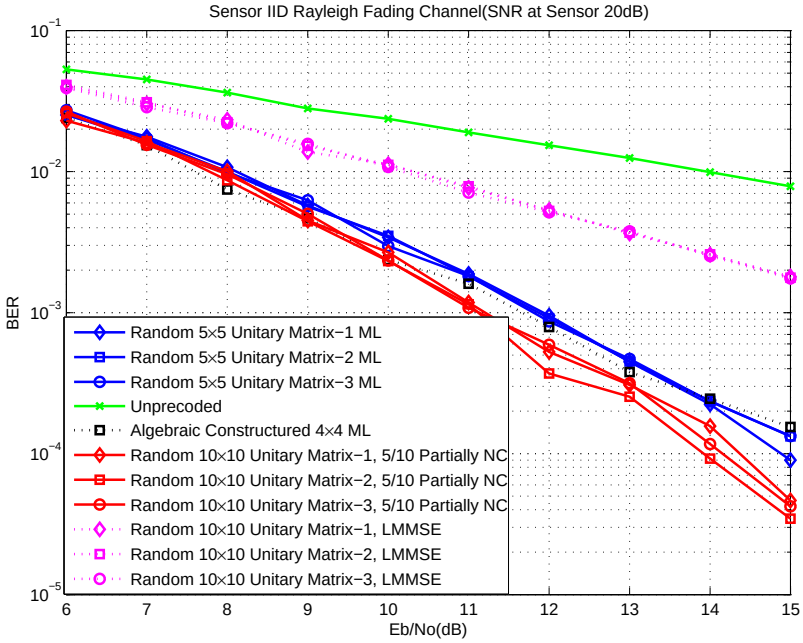

Figure 4.4: SNR at Sensor: 20dB, 5×5 ML and $5/10$ Partially Canceling and Nulling

the BER performance is not affected by the precoder; while the noise at sensor is modestly large, the approximate criteria can be obtained as in section 4.3. Though the method to construct exact optimal precoder is hard, generally the randomly generated unitary matrix is recommended to use. For the detecting scheme, the PNC method is proposed, which has good trade-off between the complexity and performance.

Chapter 5

Joint Transceiver Design towards MSE Minimization for Wireless Sensor Network

5.1 Introduction

Consider a typical wireless sensor network (WSN) comprised of a fusion center (FC) and numerous sensors that are spatially distributed and wirelessly connected to provide surveillance to the same physical event. After harvesting information from the environment, these sensors transmit distorted observations to the fusion center (FC) to perform data fusion. A central underlying problem is how to design the sensors and the fusion center to collaboratively accomplish sensing, communication and fusion task in an efficient and trust-worthy manner.

When the sensors and the fusion center are all equipped with multiple antennas and linear filters, this problem may be regarded as one of the cooperative multi-input multi-output (MIMO) beamforming design problems, which have been tackled from various perspectives [48–56]. For example [48–51] target compression (dimensionality reduction) beamforming. [48] and [49] consider the scenarios where the orthogonal multiple access channels (MAC) between the sensors and the fusion center are perfect without fading or noise. For wireless communication, the assumption of ideal channel is unrealistic and the imperfect channels are considered in [50–56]. [50] researches the problem of scalar source transmission with all sensors sharing one total transmission power and using orthogonal MAC. Imperfect coherent MAC and separate power constraint for each sensor are considered in [51], under the assumptions that all channel matrices are square and nonsingular. The work [52] and [53] are particularly relevant to our problem. [52] is the first to present a very general system model, which considers noisy and fading channels, separate power constraints and does not impose any constraints on the dimensions of beamformers or channel matrices. [52] provides the solutions to several interesting special cases of the general model for coherent MAC, such as the noiseless channel case and the no-intersymbol-interference (no-ISI) channel case. In [53], the authors develop a useful type of iterative method that is applicable to the general model in [52] for coherent MAC. All the works mentioned above take the mean square error (MSE) as performance metric. Recently, under the similar system settings of [52], joint transceiver design to maximize mutual information(MI) attract attentions and are studied in [54] and [55], with orthogonal and coherent MAC being considered respectively. The SNR maximization problem for wireless sensor network with coherent MAC is reported in [56].

It is interesting to note that the beamforming design problems in MIMO multi-sensor

decision-fusion system have significant relevance with those in other multi-agent communication networks, e.g. MIMO multi-relay and multiuser communication systems. A large number of exciting papers exist in the literature, see, for example, [58–61] and the references therein.

This chapter considers the very general coherent MAC model discussed in [52, 53]. To solve the original nonconvex joint beamforming problem, we propose several iterative optimization algorithms using the block coordinate descent (BCD) methodology, with their convergence and complexity carefully studied. Specifically our contributions include:

1) We first propose a 2 block coordinate descent (2-BCD) method that decomposes the original problem into two subproblems— one subproblem, with all the beamformers given, is a linear minimum mean square error (LMMSE) filtering problem and the other one, jointly optimizing the beamformers with the receiver given, is shown to be convex. It is worth mentioning that [52] considers the special case where the sensor-FC channels are intersymbol-interference (ISI) free (i.e. the sensor-FC channel matrix is an identity matrix) and solves the entire problem by semidefinite programming(SDP) and relaxation. Here we reformulate the joint optimization of beamformers, even with arbitrary sensor-FC channel matrices, into a second-order cone programming(SOCP) problem, which is more efficiently solvable than the general SDP problem. Convergence analysis shows that this 2-BCD algorithm guarantees its limit points to be stationary points of the original problem. Interestingly enough, although not presented in this article, the proposed 2-BCD algorithm has one more fold of importance—the convexity of its subproblem jointly optimizing beamformers can be taken advantage of by the multiplier method [70], which requires the original problem to be convex, and therefore gives birth

to decentralized solutions to the problem under the 2-BCD framework.

2) We have also attacked the MSE minimization with respect to one single beamformer and developed fully analytical solutions (possibly up to a simple one-dimension bisection search). It should be pointed out that, although the same problem has been studied in several previous papers (e.g. [53, 58, 60, 61]), we are able to carry out the analysis to the very end and thoroughly solved the problem by clearly describing the solution structure and deriving the solutions for all possible cases. Specifically, we explicitly obtain the conditions for judging the positiveness of the Lagrange multiplier. Moreover, in the zero-Lagrange-multiplier case with singular quadratic matrix, we give out the energy-preserving solution via pseudoinverse among all possible optimal solutions. To the best of our knowledge, these exact results have never been discussed in existing literature.

3) Our closed form solution for one single beamformer's update paves the way to multiple block coordinate descent algorithms. A layered-BCD algorithm is proposed, where an inner-loop cyclically optimizing each separate beamformer is embedded in the 2-BCD framework. This layered-BCD algorithm is shown to guarantee the limit points of its solution sequence to be stationary. Besides we also consider a wide class of multiple block coordinate descent algorithms with the very general essentially cyclic updating rule. It is interesting to note that this class of algorithms subsumes the one proposed in [53] as a specialized realization. Furthermore, as will be shown, by appropriately adjusting the update of each single beamformer to a proximal version and introducing approximation, the essentially cyclic multiple block coordinate descent algorithm exhibits fast converging rate, guarantees convergence to stationary points and achieves high computation efficiency.

The rest of the chapter is organized as follows: Section 5.2 introduces the system model of the joint beamforming problem in the MIMO wireless sensor network. Section 5.3 discusses the 2-BCD beamforming design approach and analyzes its convexity and convergence. Section 5.4 discusses the further decomposition of the joint optimization of beamformers, including the closed form solution to one separate beamformer's update, layered BCD algorithms, essentially cyclic BCD algorithms and their variants and convergence. Section 5.5 provides simulation verification and Section 5.6 concludes this article.

Notations: We use bold lowercase letters to denote complex vectors and bold capital letters to denote complex matrices. $\mathbf{0}$, $\mathbf{O}_{m\times n}$, and $\mathbf{I}_m$ are used to denote zero vectors, zero matrices of dimension $m \times n$, and identity matrices of order m respectively. $\mathbf{A}^T$, $\mathbf{A}^*$ and $\mathbf{A}^H$ are used to denote transpose, conjugate and conjugate transpose (Hermitian transpose) respectively of an arbitrary complex matrix $\mathbf{A}$. $\mathsf{Tr}\{\cdot\}$ denotes the trace operation of a square matrix. $|\cdot|$ denotes the modulus of a complex scalar, and $\|\cdot\|_2$ denotes the l_2-norm of a complex vector. $vec(\cdot)$ means vectorization operation of a matrix, which is performed by packing the columns of a matrix into a long one column. $\otimes$ denotes the Kronecker product. $\mathsf{diag}\{\mathbf{A}_1, \cdots, \mathbf{A}_n\}$ denotes the block diagonal matrix with its i-th diagonal block being the square complex matrix $\mathbf{A}_i$, $i \in \{1, \cdots, n\}$. $\mathsf{Re}\{x\}$ denotes the real part of a complex value x.

5.2 System Model

Consider a centralized wireless sensor network with L sensors and one fusion center where all the nodes are equipped with multiple antennae, as shown in Figure 5.1. Let

M and N_i ($i = 1, 2, \cdots, L$) be the number of antennas provisioned to the fusion center and the i-th sensor respectively. Denote $\mathbf{s}$ as the common source vector observed by all sensors. The source $\mathbf{s}$ is a complex vector of dimension K, i.e. $\mathbf{s} \in \mathbb{C}^{K\times 1}$, and is observed by all the sensors. At the i-th sensor, the source signal is linearly transformed by an observation matrix $\mathbf{K}_i \in \mathbb{C}^{J_i \times K}$ and corrupted by additive observation noise $\mathbf{n}_i$, which has zero mean and covariance matrix $\mathbf{\Sigma}_i$.

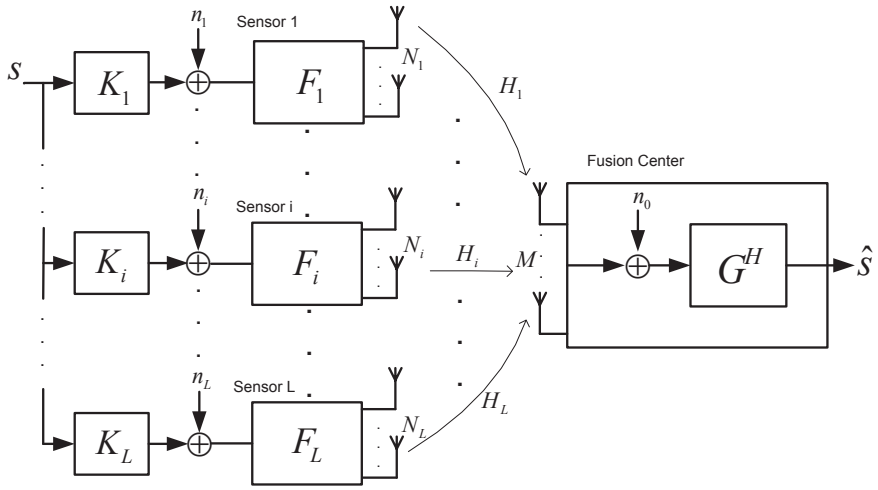

Figure 5.1: Multi-Sensor System Model

Each sensor applies some linear precoder, $\mathbf{F}_i \in \mathbb{C}^{N_i \times J_i}$, to its observation $(\mathbf{K}_i\mathbf{s}+\mathbf{n}_i)$ before sending it to the common fusion center. Denote $\mathbf{H}_i \in \mathbb{C}^{M\times N_i}$ as the fading channel between the i-th sensor and the fusion center. Here we considers the coherent MAC model, where the transmitted data is superimposed and corrupted by additive noise at the fusion center. Without loss of generality, the channel noise is modeled as a vector $\mathbf{n}_0 \in \mathbb{C}^{M\times 1}$ with zero mean and white covariance $\sigma_0^2\mathbf{I}_M$. The fusion center, after collecting all the results, applies a linear postcoder, $\mathbf{G}^H \in \mathbb{C}^{K\times M}$, to retrieve the original source $\mathbf{s}$.

This system model depicted in Figure 5.1 is the same as the general model presented

in [52, 53]. Following their convention, we assume that the system is perfectly time-synchronous (which may be realized via the GPS system) and that all the channel state information $\mathbf{H}_i$ is known (which may be achieved via channel estimation techniques). Since the sensors and the fusion center are usually distributed over a wide range of space, it is reasonable to assume that the noise $\mathbf{n}_i$ at different sensors and $\mathbf{n}_0$ at the fusion center are mutually uncorrelated.

The signal transmitted by the i-th sensor takes the form of $\mathbf{F}_i(\mathbf{K}_i\mathbf{s}+\mathbf{n}_i)$. The output $\hat{\mathbf{s}}$ of the postcoder at the fusion center is given as

$$\hat{\mathbf{s}} = \mathbf{G}^H\mathbf{r} = \mathbf{G}^H\bigg(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i(\mathbf{K}_i\mathbf{s}+\mathbf{n}_i)+\mathbf{n}_0\bigg) \tag{5.1}$$

$$= \mathbf{G}^H\bigg(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\bigg)\mathbf{s} + \mathbf{G}^H\Bigg(\underbrace{\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{n}_i+\mathbf{n}_0}_{\mathbf{n}}\Bigg), \tag{5.2}$$

where the compound noise vector $\mathbf{n}$ has covariance matrix $\mathbf{\Sigma_n}$ given by

$$\mathbf{\Sigma_n} = \sigma_0^2\mathbf{I}_M + \sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H. \tag{5.3}$$

In this chapter, we take the mean square error as a figure of merit. The mean square error matrix $\mathbf{\Phi}$ is defined as

$$\mathbf{\Phi} \triangleq \mathsf{E}\big\{\big(\mathbf{s}-\hat{\mathbf{s}}\big)\big(\mathbf{s}-\hat{\mathbf{s}}\big)^H\big\}. \tag{5.4}$$

Assume that the source signal $\mathbf{s}$ has zero mean and a covariance matrix $\mathbf{\Sigma_s} \triangleq \mathsf{E}\{\mathbf{s}\mathbf{s}^H\}$. By plugging (5.2) into (5.4), we can express the MSE matrix $\mathbf{\Phi}$ as a function of $\{\mathbf{F}_i\}$

and $\mathbf{G}$ as:

$$\begin{aligned}\mathbf{\Phi}\Big(\{\mathbf{F}_i\}_{i=1}^L,\mathbf{G}\Big)=&\mathbf{G}^H\Big(\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)\mathbf{\Sigma_s}\Big(\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)^H\mathbf{G}\\&-\mathbf{G}^H\Big(\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)\mathbf{\Sigma_s}-\mathbf{\Sigma_s}\Big(\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)^H\mathbf{G}\\&+\sum_{i=1}^L\mathbf{G}^H\mathbf{H}_i\mathbf{F}_i\mathbf{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\mathbf{G}+\sigma_0^2\mathbf{G}^H\mathbf{G}+\mathbf{\Sigma_s}.\end{aligned}\tag{5.5}$$

The total MSE is then given by

$$\mathsf{MSE}\Big(\big\{\mathbf{F}_i\big\}_{i=1}^L,\mathbf{G}\Big)\triangleq\mathsf{Tr}\Big\{\mathbf{\Phi}\Big(\big\{\mathbf{F}_i\big\}_{i=1}^L,\mathbf{G}\Big)\Big\}.\tag{5.6}$$

We consider the case where each sensor has its own transmission power constraint. This means $\mathsf{E}\{\|\mathbf{F}_i(\mathbf{K}_i\mathbf{s}+\mathbf{n}_i)\|_2^2\}=\mathsf{Tr}\{\mathbf{F}_i(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H+\mathbf{\Sigma}_i)\mathbf{F}_i^H\}\leq P_i$. The overall beamforming design problem can then be formulated as the following optimization problem:

$$(\mathsf{P}0):\min_{\{\mathbf{F}_i\}_{i=1}^L,\mathbf{G}}.\ \mathsf{MSE}\big(\{\mathbf{F}_i\}_{i=1}^L,\mathbf{G}\big),\tag{5.7a}$$

$$s.t.\ \mathsf{Tr}\big\{\mathbf{F}_i(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H+\mathbf{\Sigma}_i)\mathbf{F}_i^H\big\}\leq P_i,\ i\in\{1,\cdots,L\}.\tag{5.7b}$$

The above problem is nonconvex, which can be verified by checking the special case where $\{\mathbf{F}_i\}_{i=1}^L$ and $\mathbf{G}$ are all scalars.

The following of this chapter consults to *block coordinate descent* (BCD) method [62–65], which is also known as Gauss-Seidel method, to solve $(\mathsf{P}0)$ by partitioning the whole variables into separate groups and optimize each group (with the others being fixed) in an iterative manner. Appropriate decomposition can lead to efficiently solvable

subproblems and may also provide opportunities for parallel computation.

5.3 Two-Block Coordinate Descent (2-BCD)

In this section, we study a two block coordinate descent (2-BCD) method that decouples the design of the postcoder $\mathbf{G}$ (conditioned on the precoders), thereafter referred to as (P1), from the design of all the precoders $\{\mathbf{F}_i\}_{i=1}^L$ (conditioned on the postcoder), thereafter referred to as (P2).

5.3.1 ($P1$): Optimizing $\mathbf{G}$ given $\{\mathbf{F}_i\}$

For any given $\{\mathbf{F}_i\}_{i=1}^L$, minimizing MSE with respective to $\mathbf{G}$ becomes a strictly convex non-constrained quadratic problem (P1):

$$(\mathsf{P1}): \min_{\mathbf{G}} \mathsf{Tr}\Big\{\mathbf{\Phi}\Big(\mathbf{G}\Big|\big\{\mathbf{F}_i\big\}_{i=1}^L\Big)\Big\}. \tag{5.8}$$

By equating the derivative $\frac{\partial}{\partial \mathbf{G}^*}\mathsf{MSE}\big(\mathbf{G}\big)$ with zero, the optimal receiver is readily obtained as the well-known Wiener filter [34]

$$\mathbf{G}^{\star}_{(\mathsf{P1})}=\Big[\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)\mathbf{\Sigma_s}\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)^H+\mathbf{\Sigma_n}\Big]^{-1}\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\Big)\mathbf{\Sigma_s}, \tag{5.9}$$

where $\mathbf{\Sigma_n}$ is given in (5.3).

5.3.2 $(P2)$: Optimizing $\{\mathbf{F}_i\}$ given $\mathbf{G}$

With $\mathbf{G}$ being fixed, the subproblem (P2) minimizes MSE with respect to $\{\mathbf{F}_i\}_{i=1}^L$ is formulated as

$$(\mathsf{P2}): \min_{\{\mathbf{F}_i\}_{i=1}^L} . \ \mathsf{Tr}\Big\{\mathbf{\Phi}\Big(\big\{\mathbf{F}_i\big\}_{i=1}^L\Big|\mathbf{G}\Big)\Big\}, \tag{5.10a}$$

$$s.t. \ \mathsf{Tr}\big\{\mathbf{F}_i(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H+\mathbf{\Sigma}_i)\mathbf{F}_i^H\big\}\leq P_i, \ i\in\{1,\cdots,L\}. \tag{5.10b}$$

Below we discuss the convexity of (P2).

Theorem 5.3.1. $(P2)$ is convex with respect to $\{\mathbf{F}_i\}_{i=1}^L$.

Proof. First consider the function $f\big(\mathbf{X}\big) : \mathbb{C}^{m\times n} \mapsto \mathbb{R}, f(\mathbf{X}) = \mathsf{Tr}\{\mathbf{A}^H\mathbf{X}\mathbf{\Sigma}\mathbf{X}^H\mathbf{A}\}$, where the constant matrices $\mathbf{A}$ and $\mathbf{\Sigma}$ have appropriate dimensions and $\mathbf{\Sigma}$ is Hermitian and positive semidefinite.

By the identities $\mathsf{Tr}\{\mathbf{AB}\} = \mathsf{Tr}\{\mathbf{BA}\}$ and $\mathsf{Tr}\{\mathbf{ABCD}\} = vec^T(\mathbf{D}^T)\big[\mathbf{C}^T\otimes\mathbf{A}\big]vec(\mathbf{B})$, $f\big(\mathbf{X}\big)$ can be equivalently written as $f(\mathbf{X}) = vec^H(\mathbf{X})[\mathbf{\Sigma}^*\otimes(\mathbf{A}\mathbf{A}^H)]vec(\mathbf{X})$.

According to [67], i) $[\mathbf{A}\otimes\mathbf{B}]^H = \mathbf{A}^H\otimes\mathbf{B}^H$; ii) for any two Hermitian matrices $\mathbf{A}_{m\times m}$ and $\mathbf{B}_{n\times n}$ having eigenvalues $\{\lambda_i(\mathbf{A})\}_{i=1}^m$ and $\{\lambda_j(\mathbf{B})\}_{j=1}^n$ respectively, the eigenvalues of their Kronecker product $\mathbf{A}\otimes\mathbf{B}$ are given by $\{\lambda_i(\mathbf{A})\lambda_j(\mathbf{B})\}_{i=1,j=1}^{m,n}$. As a result, $\mathbf{A}\otimes\mathbf{B}$ is positive semidefinite when $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite.

Since $\mathbf{A}\mathbf{A}^H$ and $\mathbf{\Sigma}^*$ are both positive semidefinite, $[\mathbf{\Sigma}^*\otimes(\mathbf{A}\mathbf{A}^H)]$ is positive semidefinite and therefore $f(\mathbf{X})$ is actually a convex homogeneous quadratic function of $vec\big(\mathbf{X}\big)$.

Now substitute $\mathbf{X}$ in $f(\mathbf{X})$ by $\sum_{i=1}^L\big(\mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\big)$ and recall the fact that affine opera-

tion preserves convexity [68], the term $\mathsf{Tr}\Big\{\mathbf{G}^H\big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\big)\mathbf{\Sigma_s}\big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\mathbf{K}_i\big)^H\mathbf{G}\Big\}$ in the objective function (P2) is therefore convex with respect to $\{\mathbf{F}_i\}_{i=1}^L$. By the same reasoning, the remaining terms in the objective and the constraints of (P2) are either convex quadratic or affine functions of $\{\mathbf{F}_i\}_{i=1}^L$ and therefore the problem (P2) is convex with respective to $\{\mathbf{F}_i\}_{i=1}^L$. □

In the following we reformulate the subproblem (P2) into a standard second order cone programming(SOCP) presentation. To this end, we introduce the following notations:

$$\mathbf{f}_i \triangleq vec\big(\mathbf{F}_i\big); \quad \mathbf{g} \triangleq vec\big(\mathbf{G}\big); \tag{5.11a}$$

$$\mathbf{A}_{ij} \triangleq (\mathbf{K}_j\mathbf{\Sigma_s}\mathbf{K}_i^H)^T \otimes \Big(\mathbf{H}_i^H\mathbf{G}\mathbf{G}^H\mathbf{H}_j\Big); \tag{5.11b}$$

$$\mathbf{B}_i \triangleq (\mathbf{K}_i\mathbf{\Sigma_s})^T \otimes \mathbf{H}_i; \tag{5.11c}$$

$$\mathbf{C}_i \triangleq \mathbf{\Sigma}_i^* \otimes \Big(\mathbf{H}_i^H\mathbf{G}\mathbf{G}^H\mathbf{H}_i\Big). \tag{5.11d}$$

By the identity $\mathsf{Tr}\{\mathbf{ABCD}\} = vec^T(\mathbf{D}^T)\big[\mathbf{C}^T \otimes \mathbf{A}\big]vec(\mathbf{B})$ and the above notations, we can rewrite the MSE in (P2) as

$$\begin{aligned}\mathsf{MSE}\Big(\big\{\mathbf{f}_i\big\}_{i=1}^L\Big|\mathbf{g}\Big) = &\sum_{i=1}^L\sum_{j=1}^L \mathbf{f}_i^H\mathbf{A}_{ij}\mathbf{f}_j - 2\mathsf{Re}\Big(\sum_{i=1}^L \mathbf{g}^H\mathbf{B}_i\mathbf{f}_i\Big)\\ &+ \sum_{i=1}^L \mathbf{f}_i^H\mathbf{C}_i\mathbf{f}_i + \sigma_0^2\|\mathbf{g}\|^2 + \mathsf{Tr}\big\{\mathbf{\Sigma_s}\big\}.\end{aligned} \tag{5.12}$$

By further denoting

$$\mathbf{f}^T \triangleq \left[\mathbf{f}_1^T, \cdots, \mathbf{f}_i^T, \cdots, \mathbf{f}_L^T\right]; \tag{5.13a}$$

$$\mathbf{A} \triangleq \begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{A}_{1,2} & \cdots & \mathbf{A}_{1,L} \\ \mathbf{A}_{2,1} & \mathbf{A}_{2,2} & \cdots & \mathbf{A}_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{L,1} & \mathbf{A}_{L,2} & \cdots & \mathbf{A}_{L,L} \end{bmatrix}; \tag{5.13b}$$

$$\mathbf{B} \triangleq \Big[\mathbf{B}_1, \cdots, \mathbf{B}_i, \cdots, \mathbf{B}_L\Big]; \tag{5.13c}$$

$$\mathbf{C} \triangleq \mathsf{diag}\Big\{\mathbf{C}_1, \cdots, \mathbf{C}_i, \cdots, \mathbf{C}_L\Big\}; \tag{5.13d}$$

$$\mathbf{D}_i \triangleq \mathsf{diag}\Big\{\mathbf{O}_{\sum_{j=1}^{i-1} J_j N_j}, \mathbf{E}_i, \mathbf{O}_{\sum_{j=i+1}^{L} J_j N_j}\Big\}, i \in \{1, \cdots, L\}; \tag{5.13e}$$

$$\mathbf{E}_i \triangleq \big(\mathbf{K}_i \mathbf{\Sigma_s} \mathbf{K}_i^H + \mathbf{\Sigma}_i\big)^T \otimes \mathbf{I}_{N_i}, \quad i \in \{1, \cdots, L\}; \tag{5.13f}$$

$$c \triangleq \mathsf{Tr}\{\mathbf{\Sigma_s}\} + \sigma_0^2 \|\mathbf{g}\|^2, \tag{5.13g}$$

the problem (P2) can be rewritten as (P2′):

$$(\mathsf{P2}') : \min_{\mathbf{f}} \mathbf{f}^H \big(\mathbf{A}+\mathbf{C}\big)\mathbf{f} - 2\mathsf{Re}\{\mathbf{g}^H \mathbf{B}\mathbf{f}\} + c, \tag{5.14a}$$

$$s.t.\ \mathbf{f}^H \mathbf{D}_i \mathbf{f} \leq P_i, \quad i \in \{1, \cdots, L\}. \tag{5.14b}$$

As proved by Theorem 5.3.1, (P2′) (or equivalently (P2)) is convex, which implies $(\mathbf{A}+\mathbf{C})$ is positive semidefinite. Thus the square root $(\mathbf{A}+\mathbf{C})^{\frac{1}{2}}$ exists. The above

problem can therefore be reformulated in an SOCP form as follows

$$(\mathsf{P}2_{SOCP}): \min_{\mathbf{f},t,s} . \ t, \tag{5.15a}$$

$$\text{s.t. } s - 2\mathsf{Re}\{\mathbf{g}^H\mathbf{B}\mathbf{f}\} + c \leq t; \tag{5.15b}$$

$$\left\| \begin{array}{c} (\mathbf{A}+\mathbf{C})^{\frac{1}{2}}\mathbf{f} \\ \frac{s-1}{2} \end{array} \right\|_2 \leq \frac{s+1}{2}; \tag{5.15c}$$

$$\left\| \begin{array}{c} \mathbf{D}_i^{\frac{1}{2}}\mathbf{f} \\ \frac{P_i-1}{2} \end{array} \right\|_2 \leq \frac{P_i+1}{2}, \ \ i \in \{1, \cdots, L\}; \tag{5.15d}$$

$(\mathsf{P}2_{SOCP})$ can be numerically solved by off-the-shelf convex programming solvers, such as CVX [69].

Summarizing the above discussions, the problem $(\mathsf{P}0)$ can be solved by a 2-BCD algorithm: updating $\mathbf{G}$ by solving $(\mathsf{P}1)$ and updating $\left\{\mathbf{F}_i\right\}_{i=1}^{L}$ by solving $(\mathsf{P}2')$ alternatively, which is summarized in Algorithm 1.

Algorithm 1: 2-BCD Algorithm to Solve $(\mathsf{P}0)$

1 **Initialization**: Randomly generate feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^{L}$, $i \in \{1, \cdots, L\}$; Compute $\mathbf{G}^{(0)}$ using (5.9);
2 **repeat**
3 With $\mathbf{G}^{(j-1)}$ fixed, solve $(\mathsf{P}2')$ and obtain $\{\mathbf{F}_i^{(j)}\}_{i=1}^{L}$;
4 With $\{\mathbf{F}_i^{(j)}\}_{i=1}^{L}$ fixed, compute $\mathbf{G}^{(j)}$ using (5.9);
5 **until** *decrease of* MSE *is small enough or predefined number of iterations is reached*;

5.3.3 Convergence of 2-BCD Algorithm

In this subsection we study the convergence of the above 2-BCD algorithm. Consider the optimization problem $\min\{f(\mathbf{x})|\mathbf{x} \in \mathcal{X}\}$ with $f(\cdot)$ being continuously differentiable and the feasible domain $\mathcal{X}$ being closed and nonempty. A point $\mathbf{x}_0 \in \mathcal{X}$ is a *stationary point* if and only if $\nabla f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \geq 0$, $\forall \mathbf{x} \in \mathcal{X}$, where $\nabla f(\mathbf{x}_0)$ denotes the gradient of f at $\mathbf{x}_0$. For the proposed 2-BCD algorithm, we have the following convergence conclusion.

Theorem 5.3.2. The objective sequence $\{\mathsf{MSE}^{(j)}\}_{j=0}^{\infty}$ generated by the 2-BCD algorithm in Algorithm 1 is monotonically decreasing. If $\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H \succ 0$ or $\mathbf{\Sigma}_i \succ 0$ for all $i \in \{1, \cdots, L\}$, the solution sequence $\left\{\{\mathbf{F}_i^{(j)}\}_{i=1}^{L}, \mathbf{G}^{(j)}\right\}_{j=1}^{\infty}$ generated by the 2-BCD algorithm has limit points and each limit point of $\left\{\{\mathbf{F}_i^{(j)}\}_{i=1}^{L}, \mathbf{G}^{(j)}\right\}_{j=1}^{\infty}$ is a stationary point of $(P0)$.

Proof. Since each block update solves a minimization problem, MSE keeps decreasing. Let $\mathcal{X}_i = \left\{\mathbf{X} \in \mathbb{C}^{N_i \times J_i} \middle| \mathsf{Tr}\{\mathbf{X}(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H + \mathbf{\Sigma}_i)\mathbf{X}^H\} \leq P_i\right\}$, for $i = 1, \cdots, L$ and $\mathcal{X}_{L+1} = \mathbb{C}^{M \times K}$. Under the strictly positive definiteness assumption of $\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H$ or $\mathbf{\Sigma}_i$, we have $\left(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H + \mathbf{\Sigma}_i\right) \succ 0$ and thus $\left(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H + \mathbf{\Sigma}_i\right)^T \otimes \mathbf{I}_{N_i} \succ 0$ for all $i \in \{1, \cdots, L\}$. This implies that the null space of $\left(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H + \mathbf{\Sigma}_i\right)^T \otimes \mathbf{I}_{N_i}$ is $\{\mathbf{0}\}$ and consequently $\mathbf{f}_i$ has to be bounded to satisfy power constraint. Therefore $\mathcal{X}_i$ is bounded for all $i \in \{1, \cdots, L\}$. Since the feasible set for each $\mathbf{F}_i$ is bounded, by Bolzano-Weierstrass theorem, there exists a convergent subsequence $\left\{\{\mathbf{F}_i^{(j_k)}\}_{i=1}^{L}\right\}_{k=1}^{\infty}$. Since $\mathbf{G}$ is updated by equation (5.9) as a continuous function of $\{\mathbf{F}_i\}_{i=1}^{L}$, the subsequence $\{\mathbf{G}^{(j_k+1)}\}_{k=1}^{\infty}$ also converges and thus bounded. By further restricting to a subsequence of $\left\{\{\mathbf{F}_i^{(j_k+1)}\}, \mathbf{G}^{(j_k+1)}\right\}_{k=1}^{\infty}$, we can obtain a convergent subsequence of $\left\{\{\mathbf{F}_i^{(j)}\}_{i=1}^{L}, \mathbf{G}^{(j)}\right\}_{j=1}^{\infty}$.

Since Algorithm 1 is a two block coordinate descent procedure and the problem (P0) has continuously differentiable objective and closed and convex feasible domain, Corollary 2 in [64] is valid to invoke, we conclude that any limit point of $\left\{\{\mathbf{F}_i^{(j)}\}_{i=1}^L, \mathbf{G}^{(j)}\right\}_{j=1}^{\infty}$ is a stationary point of (P0). □

5.4 Multi-Block Coordinate Descent

For the above 2-BCD algorithm, although we can solve the subproblem (P2) as a standard SOCP problem, its closed-form solution is still inaccessible. The complexity for solving (P2) can be shown to be $\mathcal{O}\left(\sqrt{L}\left(\sum_{i=1}^L N_i J_i\right)^3\right)$, This implies that when the sensor network under consideration has a large number of sensors and/or antennae, the complexity for solving (P2) can be rather daunting. This motivates us to search for more efficient ways to update sensor's beamformer.

5.4.1 Further Decoupling of ($P2$) and Closed-Form Solution

Looking back to problem (P2), although it has separable power constraints, its quadratic terms in its objective tangles different sensors' beamformers together and thus makes the Karush-Kuhn-Tucker(KKT) conditions of (P2) analytically unsolvable. Here we adopt the BCD methodology to further decompose the subproblem (P2). Instead of optimizing all the $\mathbf{F}_i$'s in a single batch, we optimize one $\mathbf{f}_i$ at a time with the others being fixed. By introducing the notation $\mathbf{q}_i \triangleq \sum_{j=1,j\neq i}^L \mathbf{A}_{ij}\mathbf{f}_j$, each subproblem (P2$_i'$)

of (P2′) is given as

$$(\mathsf{P2}_i'): \min_{\mathbf{f}_i} \mathbf{f}_i^H \big(\mathbf{A}_{ii}+\mathbf{C}_i\big)\mathbf{f}_i + 2\mathsf{Re}\{\mathbf{q}_i^H\mathbf{f}_i\}-2\mathsf{Re}\{\mathbf{g}^H\mathbf{B}_i\mathbf{f}_i\} \tag{5.16a}$$

$$s.t.\ \mathbf{f}_i^H\mathbf{E}_i\mathbf{f}_i \leq P_i. \tag{5.16b}$$

Now our problem boils down to solving the simpler problem (P2′$_i$), for $i = 1, \cdots, L$. The following theorem provides an *almost* closed-form solution to (P2′$_i$). The only reason that this is not a *fully* closed-form solution is because it may involve a bisection search to determine the value of a positive real number.

Theorem 5.4.1. Assume $\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H \succ 0$ or $\mathbf{\Sigma}_i \succ 0$. Define parameters $\mathbf{M}_i$, $\mathbf{U}_i$ and $\mathbf{p}_i$ as in equations (5.26) in the appendix, r_i as the rank of $\mathbf{M}_i$ and $p_{i,k}$ as the i-th entry of $\mathbf{p}_i$. The solution to ($P2_i'$) is given as follows:

CASE (I)—if either of the following two conditions holds:

i) $\exists k \in \{r_i+1, \cdots, J_iN_i\}$ such that $|p_{i,k}| \neq 0$;

or ii) $\sum_{k=r_i+1}^{J_iN_i} |p_{i,k}| = 0$ and $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} > P_i$.

The optimal solution to ($P2_i'$) is given by

$$\mathbf{f}_i^\star = \big(\mathbf{A}_{ii}+\mathbf{C}_i+\mu_i^\star\mathbf{E}_i\big)^{-1}\big(\mathbf{B}_i^H\mathbf{g}-\mathbf{q}_i\big), \tag{5.17}$$

with the positive value $\mu_i^\star$ being the unique solution to the equation: $g_i(\mu_i) = \sum_{k=1}^{J_iN_i} \frac{|p_{i,k}|^2}{(\lambda_{i,k}+\mu_i)^2} = P_i$. An interval $\big[lbd_i, ubd_i\big]$ containing $\mu_i^\star$ is determined by Lemma 5.4.1 which comes later.

CASE (II)—$\sum_{k=r_i+1}^{J_iN_i} |p_{i,k}| = 0$ and $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} \leq P_i$,

The optimal solution to $(P2_i')$ is given by

$$\mathbf{f}_i^\star = \mathbf{E}_i^{-\frac{1}{2}}\Big(\mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{A}_{ii}+\mathbf{C}_i\big)\mathbf{E}_i^{-\frac{1}{2}}\Big)^\dagger \mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{B}_i^H\mathbf{g}-\mathbf{q}_i\big). \tag{5.18}$$

Proof. See Appendix 5.7.1. □

Here we have several comments and supplementary discussions on the solution to $(\mathsf{P}2_i')$.

Comment 5.4.1. *When $\mu_i^\star = 0$ and $\mathbf{M}_i$ is singular, the solution to ($\mathsf{P}2_i'$) is usually non-unique. According to the proof procedure in Appendix 5.7.1, (5.18) is actually the power-preserving optimal solution, which has the minimal transmission power among all optimal solutions to ($\mathsf{P}2_i'$).*

Comment 5.4.2. *It is worth noting that the three cases discussed in the proof of Theorem 5.4.1,* CASE*(I)-case i),* CASE*(I)-case ii) and* CASE*(II), are mutually exclusive events. One and only one case will occur.*

Comment 5.4.3. *The problem of minimizing MSE with respect to one separate beamformer with one power constraint is a rather standard problem that has been discussed in previous works such as [53, 58, 60, 61]. A big contribution here is that we have* fully *solved this problem by clearly identifying the solution structure and writing out the almost closed-form solutions for all possible cases, whereas the previous papers have not. One key consideration is the case of rank deficient $\mathbf{M}_i$ for zero $\mu_i^\star$. Although [58] and [53] mention that $\mu_i^\star$ can be zero, the solution for singular $\mathbf{M}_i$ in this case is missing. In fact when $\mathbf{M}_i$ does not have full rank and $\mu_i^\star$ is zero, its inverse does not exist and consequently the solutions given in [53, 58, 60, 61] do not stand any more (they all provide solutions by matrix inversion). It is noted that [53] imposes more assumptions*

on the number of antennas to exclude some cases where $\mathbf{M}_i$ *is rank deficient. However these assumptions undermine the generality of the system model and, still, adverse channel parameters can result in rank deficiency of* $\mathbf{M}_i$*. Turns out, the rank deficiency scenario is actually not rare. In fact, whenever* $K < N_i$ *or* $M < N_i$ *holds, the matrices* $\mathbf{A}_{ii}$ *and* $\mathbf{C}_i$ *are both born rank deficient. If they share common nonzero components of null space,* $\mathbf{M}_i$ *will be rank deficient. For example, consider the simple case where* $\mathbf{K}_i = \mathbf{I}_K$*,* $\mathbf{\Sigma_s} = \sigma_s^2\mathbf{I}$*,* $\mathbf{\Sigma}_i = \sigma_i^2\mathbf{I}$ *and* $\min(K, M) < N_i$*. At this time* $\mathbf{M}_i$ *is not of full rank. Besides inappropriate channel parameters* $\mathbf{H}_i$ *can also generate rank deficient* $\mathbf{M}_i$*. Thus taking the rank deficiency of* $\mathbf{M}_i$ *when* $\mu_\star = 0$ *into consideration is both necessary and meaningful.*

Comment 5.4.4. *In the special case where* $K = J_i = 1$*, the fully closed form solution to (*$\mathsf{P2}_i$*) does exist! At this time, the optimal* $\mu_i^\star$ *and* $\mathbf{f}_i^\star$ *can be obtained analytically without bisection search. In this case, eigenvalue decomposition is also unnecessary. So when* $K = J_i = 1$*, solving (*$\mathsf{P2}_i$*) is extremely efficient. The details can be found in [57]*[1]*.*

Recall that in CASE (I) of Thoerem 5.4.1, $\mu_i^\star$ is obtained as the solution to $g_i(\mu_i) = P_i$. This equation generally has no analytic solution. Fortunately $g_i(\mu_i)$ is strictly decreasing in μ_i and thus the equation can be efficiently solved by a bisection search. The following lemma provides an interval $[lbd_i, ubd_i]$ containing the positive $\mu_i^\star$, from which the bisection search to determine $\mu_i^\star$ can be started.

Lemma 5.4.1. The positive $\mu_i^\star$ in $(P2_i')$ (i.e. CASE (I) in Theorem 5.4.1) has the following lower bound lbd_i and upper bound ubd_i:

[1] [57] actually solves an approximation of problem ($\mathsf{P2}_i'$) with scalared source , where a specific affine term of $\mathbf{f}_i$ in the objective of ($\mathsf{P2}_i'$) is approximated by its latest value (approximation is discussed in subsection 5.4.4 of this chapter). However fully analytic solution of ($\mathsf{P2}_i$) can be obtained by following very similar lines as [57] without introducing approximation of $\mathbf{f}_i$.

i) For subcase i)

$$lbd_i = \left[\frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,1}\right]^+, \quad ubd_i = \frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}}; \tag{5.19}$$

ii) For subcase ii)

$$lbd_i = \left[\frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,1}\right]^+, \quad ubd_i = \frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,r_i}, \tag{5.20}$$

where $[x]^+ = \max\{0, x\}$.

Proof. For subcase i), by definition of $g_i(\mu_i)$ in (5.30), we have

$$\begin{aligned}\frac{\|\mathbf{p}_i\|_2^2}{(\mu_i + \lambda_{i,1})^2} = \frac{\sum_{k=1}^{J_i N_i} |p_{i,k}|^2}{(\mu_i + \lambda_{i,1})^2} \leq g_i(\mu_i) = P_i \\ \leq \frac{\sum_{k=1}^{J_i N_i} |p_{i,k}|^2}{\mu_i^2} = \frac{\|\mathbf{p}_i\|_2^2}{\mu_i^2},\end{aligned} \tag{5.21}$$

which can be equivalently written as

$$\frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,1} \leq \mu_i \leq \frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}}. \tag{5.22}$$

Also notice that $\mu_i^\star$ should be positive; the bounds in (5.19) thus follow.

For subcase ii), by assumption, $\sum_{k=r_i+1}^{J_i N_i} |p_{i,k}|^2 = 0$. This leads to

$$\begin{aligned}\frac{\|\mathbf{p}_i\|_2^2}{(\mu_i + \lambda_{i,1})^2} = \frac{\sum_{k=1}^{r_i} |p_{i,k}|^2}{(\mu_i + \lambda_{i,1})^2} \leq g_i(\mu_i) = P_i \\ \leq \frac{\sum_{k=1}^{r_i} |p_{i,k}|^2}{(\mu_i + \lambda_{i,r_i})^2} = \frac{\|\mathbf{p}_i\|_2^2}{(\mu_i + \lambda_{i,r_i})^2}.\end{aligned} \tag{5.23}$$

Following the same line of derivation as in subcase i), we obtain the bounds in (5.20).

□

Algorithm 2: Solving the Problem ($\mathsf{P}2_i'$)

Initialization: Perform eigenvalue decomposition $\mathbf{M}_i = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^H$; Calculate $\mathbf{p}_i$ using (5.26d);

if $\Big(\exists k \in \{r_i+1, \cdots, J_iN_i\}$ *s.t.* $|p_{i,k}| \neq 0\Big)$ or $\Big(\sum_{k=r_i+1}^{J_iN_i} |p_{i,k}|^2 = 0$ *and* $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} > P_i\Big)$ **then**

  Determine bounds lbd_i and ubd_i via (5.19) or (5.20) ;

  Bisection search on $\big[lbd_i, ubd_i\big]$ to determine $\mu_i^\star$;

  $\mathbf{f}_i^\star = \big(\mathbf{A}_{ii} + \mathbf{C}_i + \mu_i^\star \mathbf{E}_i\big)^{-1}\big(\mathbf{B}_i^H\mathbf{g} - \mathbf{q}_i\big)$;

else

  $\mathbf{f}_i^\star = \mathbf{E}_i^{-\frac{1}{2}}\Big(\mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{A}_{ii} + \mathbf{C}_i\big)\mathbf{E}_i^{-\frac{1}{2}}\Big)^\dagger \mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{B}_i^H\mathbf{g} - \mathbf{q}_i\big)$;

end

Algorithm 2 summarizes the results obtained in Theorem 5.4.1 and Lemma 5.4.1 and provides a (nearly) closed-form solution to ($\mathsf{P}2_i'$).

5.4.2 Layered-BCD Algorithm

The above analysis of ($\mathsf{P}2_i'$), combined with ($\mathsf{P}1$), naturally leads to a nested or layered-BCD algorithm, that can be used to analytically solve the joint beamforming problem ($\mathsf{P}0$). The algorithm consists of two loops (two layers). The outer-loop is a two-block descent procedure alternatively optimizing $\mathbf{G}$ and $\{\mathbf{F}_i\}_{i=1}^L$, and the inner-loop further decomposes the optimization of $\{\mathbf{F}_i\}_{i=1}^L$ into an L-block descent procedure operated in an iterative round robin fashion. Algorithm 3 outlines the overall procedure. As will be seen in the next, this layered-BCD has strong convergence property.

Theorem 5.4.2. Assume that $\mathbf{K}_i\mathbf{\Sigma}_\mathbf{s}\mathbf{K}_i^H \succ 0$ or $\mathbf{\Sigma}_i \succ 0$, $\forall i \in \{1, \cdots, L\}$. The objective sequence $\{\mathsf{MSE}^{(j)}\}_{j=0}^\infty$ generated by Algorithm 3 is monotonically decreasing. The

Algorithm 3: Layered-BCD Algorithm to Solve (P0)

Initialization: Randomly generate feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^{L}$;
Obtain $\mathbf{G}^{(0)}$ by (5.9);
repeat
  repeat
    for $i = 1;\ i <= L;\ i++$ **do**
      Given $\mathbf{G}$ and $\{\mathbf{F}_j\}_{j\neq i}$, update $\mathbf{F}_i$ by Theorem 5.4.1;
    end
  until *decrease of* MSE *is sufficiently small*;
  Given $\{\mathbf{F}_i\}_{i=1}^{L}$, update $\mathbf{G}$ via (5.9) ;
until *decrease of* MSE *is sufficiently small or predefined number of iterations is reached*;

solution sequence $\left\{\{\mathbf{F}_i^{(j)}\}_{i=1}^{L}, \mathbf{G}^{(j)}\right\}_{j=1}^{\infty}$ generated by Algorithm 3 has limit points, and each limit point is a stationary point of $(P0)$.

Proof. The proof of the monotonicity of $\{\mathsf{MSE}^{(j)}\}_{j=0}^{\infty}$ and the existence of limit points for the solution sequence follows the same lines as those of Theorem 5.3.2.

From Theorem 5.3.1, given $\mathbf{G}$, the objective function $\mathsf{MSE}\big(\{\mathbf{F}_i\}_{i=1}^{L}\big|\mathbf{G}\big)$ of Problem (5.14) is convex (and therefore, of course, pseudoconvex) with respect to $\{\mathbf{f}_i\}_{i=1}^{L}$. Since the objective $\mathsf{MSE}\big(\{\mathbf{F}_i\}_{i=1}^{L}\big|\mathbf{G}\big)$ in (P2) is continuous and the feasible domain of $\{\mathbf{F}_i\}_{i=1}^{L}$ is bounded, there exists some feasible point $\{\bar{\mathbf{F}}_i\}_{i=1}^{L}$ making the level set $\Big\{\{\mathbf{F}_i\}_{i=1}^{L} \in \mathbb{C}^{J_1N_1\times 1}\times\cdots\times\mathbb{C}^{J_LN_L\times 1}\Big|\ \mathsf{MSE}\big(\{\mathbf{F}_i\}_{i=1}^{L}\big|\mathbf{G}\big) \leq \mathsf{MSE}\big(\{\bar{\mathbf{F}}_i\}_{i=1}^{L}\big|\mathbf{G}\big)\Big\}$ closed and bounded. Thus Proposition 6 in [64] is valid to invoke. For a given $\mathbf{G}$ at any step of outer-loop, the inner loop generates limit point(s) converging to a stationary point of the problem (P2). Since (P2) is a convex problem, any stationary point is actually an optimal solution [63]. Therefore the subproblem (P2) is actually globally solved. By Theorem 5.3.2, each limit point of solution sequence is a stationary point of the original problem (P0). □

Although the convergence analysis in Theorem 5.4.2 states that the layered-BCD algorithm guarantees convergence, it requires the inner-loop to iterate numerous times to converge sufficiently. In fact if each inner loop is performed with a small number of iterations, the layered BCD algorithm becomes a specialized essentially cyclic BCD algorithm, which will be discussed in next subsection.

5.4.3 Essentially Cyclic $(L+1)$-BCD Algorithm

In this subsection, we propose an $(L+1)$-BCD algorithm, where in each update the linear FC receiver or one single beamformer is updated efficiently by equation (5.9) or Theorem 5.4.1 respectively. Compared to the 2-BCD algorithm, the block updating rule for multiple block coordinate descent method can have various patterns. Here we adopt a very general updating manner called *essentially cyclic rule* [65]. For essentially cyclic update rule, there exists a positive integer T, which is called period, such that each block of variables is updated at least once within any consecutive T updates. The classical Gauss-Seidel method is actually a special case of essentially cyclic rule with its period T being exactly the number of blocks of variables.

For the convergence of essentially cyclic BCD algorithm, when the whole solution sequence converges, the limit of the solution sequence is stationary. In fact, assume that the sequence $\left\{\{\mathbf{F}_i^{(j)}\}_{i=1}^L, \mathbf{G}^{(j)}\right\}_{j=1}^{\infty}$ converges to the limit point $\bar{\mathbf{X}} \triangleq \left\{\{\bar{\mathbf{F}}_i\}_{i=1}^L, \bar{\mathbf{G}}\right\}$. Denote $\mathbf{X} = \left\{\{\mathbf{F}_i\}_{i=1}^L, \mathbf{G}\right\}$ and $\mathbf{X}_i$ as the i-th block of $\mathbf{X}$, which can be $\mathbf{G}$ or $\mathbf{F}_j$, $\forall j \in \{1, \cdots, L\}$, and $\mathbf{X}_{\bar{i}}$ as the variables other than $\mathbf{X}_i$, i.e. $\mathbf{X}_{\bar{i}} = \{\mathbf{X}\}\backslash\{\mathbf{X}_i\}$. Since $\mathbf{X}_i^{(j+1)}$ minimizes MSE with given $\{\mathbf{X}_{\bar{i}}^{(j)}\}$, as optimality conditions, we have $\mathsf{Tr}\{\nabla_{\mathbf{X}_i}\mathsf{MSE}(\mathbf{X}_i^{(j+1)}, \mathbf{X}_{\bar{i}}^{(j)})^T\left(\mathbf{X}_i - \mathbf{X}_i^{(j+1)}\right)\} \geq 0$ for any feasible $\mathbf{X}_i$. Since $\{\mathbf{X}_{\bar{i}}^{(j)}\} \rightarrow$

$\bar{\mathbf{X}}_{\bar{i}}, \mathbf{X}_i^{(j+1)} \to \bar{\mathbf{X}}_i$ and MSE is continuously differentiable, we have $\mathsf{Tr}\{\nabla_{\mathbf{X}_i}\mathsf{MSE}(\bar{\mathbf{X}})^T\big(\mathbf{X}-\bar{\mathbf{X}}_i\big)\} \geq 0$ for any feasible $\mathbf{X}_i$, $\forall i \in \{1, \cdots, L+1\}$. By summing up all $L+1$ variable blocks, we obtain $\mathsf{Tr}\{\nabla_{\mathbf{X}}\mathsf{MSE}(\bar{\mathbf{X}})^T\big(\mathbf{X}-\bar{\mathbf{X}}\big)\} \geq 0$ for any feasible $\mathbf{X}$. This suggests that the convergent limit point $\big\{\{\bar{\mathbf{F}}_i\}_{i=1}^L, \bar{\mathbf{G}}\big\}$ is actually a stationary point of (P0).

However the assumption that the whole solution sequence converges is actually a very strong assumption and cannot be theoretically proved, although extensive numerical results show that this fact seem always hold in practice for our problem.

For rigorous proof of the convergence to stationary points of BCD algorithms, one usually requires uniqueness of solutions for each block update, as the analysis performed in [63–65]. Without the uniqueness assumptions, convergence to stationary points is not guaranteed and a counter example has been reported in [66], where the solution sequence is always far away from stationary points. In retrospect to Theorem 5.4.1, specific parameter settings (CASE(II) with singular $(\mathbf{A}_{ii}+\mathbf{C}_i)$ and zero $\mu_i^\star$) will result in infinitely many optimal solutions to ($\mathsf{P}2_i'$). To overcome this difficulty, we adopt proximal method (Exercise 2.7.1 in [63]), which locally modifies the ($\mathsf{P}2_i'$) by imposing a squared norm and guarantees that each block update is uniquely solved.

Specifically, to update the i-th beamformer, we consult to the proximal version objective $\mathsf{MSE}\big(\mathbf{f}_i\big|\{\mathbf{f}_j\}_{j\neq i}, \mathbf{g}\big)+\kappa\|\mathbf{f}_i-\hat{\mathbf{f}}_i\|_2^2$ of ($\mathsf{P}2_i'$) with $\hat{\mathbf{f}}_i$ being the latest value of $\mathbf{f}_i$ until the current update and κ being any positive real constant. Thus the problem updating the i-th sensor's beamformer is equivalent to ($\mathsf{P}4_i$) as follows

$$\begin{aligned}(\mathsf{P}4_i): \min_{\mathbf{f}_i}\ & \mathbf{f}_i^H\big(\mathbf{A}_{ii}+\mathbf{C}_i+\kappa\mathbf{I}_{N_iJ_i}\big)\mathbf{f}_i+2\mathsf{Re}\big\{\big(\mathbf{q}_i^H-\mathbf{g}^H\mathbf{B}_i-\kappa\hat{\mathbf{f}}_i^H\big)\mathbf{f}_i\big\} \\ s.t.\ & \mathbf{f}_i^H\mathbf{E}_i\mathbf{f}_i \leq P_i. \end{aligned} \tag{5.24}$$

As shown by the following theorem, the proximal version of any essentially cyclic $(L+1)$-BCD algorithm guarantees monotonic decreasing of objective and stationary-point-achieving convergence of the solution sequence.

Theorem 5.4.3. Assume that $\mathbf{K}_i\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{K}_i^H \succ 0$ or $\boldsymbol{\Sigma}_i \succ 0, \forall i \in \{1, \cdots, L\}$. By updating $\mathbf{G}$ and $\mathbf{F}_i$ by solving $(P1)$ and $(P4_i)$ respectively, any essentially cyclic $(L+1)$-BCD algorithm generates monotonically decreasing MSE sequence and the solution sequence has limit points with each limit point being a stationary point of the original problem $(P0)$.

Proof. See Appendix 5.7.2. □

Note the solution to $(\mathsf{P}4_i)$ can be easily obtained by Theorem 3 with the terms $\left(\mathbf{A}_{ii}+\mathbf{C}_i\right)$ and $\left(\mathbf{B}_i^H\mathbf{g}–\mathbf{q}_i\right)$ being replaced by $\left(\mathbf{A}_{ii}+\mathbf{C}_i+\kappa\mathbf{I}_{N_iJ_i}\right)$ and $\left(\mathbf{B}_i^H\mathbf{g}–\mathbf{q}_i+\kappa\hat{\mathbf{f}}_i\right)$ respectively and no additional complexity is required.

Recall the layered-BCD algorithm discussed in previous subsection, when the inner-loop is performed by small number of iterations, it actually reduces to a specialized essentially cyclic BCD algorithm. One special case is the iterative algorithm proposed in [53] whose inner-loop updates each beamformer for once. According to the above theorem, by updating each beamformer with the proximal method, the convergence to stationary points can be guaranteed.

One drawback of the above proximal update is its slow convergence rate, as will be shown in Section 5.5. However this shortcoming can be well compensated by the following acceleration scheme in the next subsection.

5.4.4 Acceleration by Approximation

The aforementioned $(L+1)$-BCD algorithm can be accelerated by introducing approximation when updating single beamformer $\mathbf{F}_i$ in $(\mathsf{P}2_i')$. In addition to setting the $\{\mathbf{F}_j\}_{j\neq i}$ as known and fixed, we assume that the term $\mathbf{A}_{ii}\mathbf{f}_i$ is also known by leveraging the value of $\mathbf{f}_i$ in the previous updates. In other words, we define $\hat{\mathbf{q}}_i = \sum_{j=1,j\neq i}^{L} \mathbf{A}_{ij}\mathbf{f}_j + \mathbf{A}_{ii}\hat{\mathbf{f}}_i = \mathbf{q}_i + \mathbf{A}_{ii}\hat{\mathbf{f}}_i$ with $\hat{\mathbf{f}}_i$ being the latest value of $\mathbf{f}_i$. Thus to update $\mathbf{f}_i$ we solve the approximate version $(\mathsf{P}5_i)$ of $(\mathsf{P}2_i')$ as follows

$$(\mathsf{P}5_i): \min_{\mathbf{f}_i} \mathbf{f}_i^H \mathbf{C}_i \mathbf{f}_i + 2\mathsf{Re}\{\hat{\mathbf{q}}_i^H \mathbf{f}_i\} - 2\mathsf{Re}\{\mathbf{g}^H \mathbf{B}_i \mathbf{f}_i\} \tag{5.25a}$$

$$s.t.\ \mathbf{f}_i^H \mathbf{E}_i \mathbf{f}_i \leq P_i. \tag{5.25b}$$

The problem $(\mathsf{P}5_i)$ can still be efficiently solved by Theorem 5.4.1. Interestingly enough, this approximation can significantly improve the convergence rate of the cyclic-BCD procedure!

Actually similar idea appears in [58], where the precoders of multiusers is updated in a cyclic manner. In Implementation 2 (Table II) of [58], with others being fixed, one separate precoder is updated by minimizing the total MSE function with some terms of the to-be-updated precoder approximated by previous values. As reflected by the extensive numerical results in [58], this approximated BCD implementation has surprisingly faster convergence compared to the original one (Implementation I in Table I) in [58].

The surprisingly fast convergence of the approximate update inspires us the idea that it can become perfect complement of the aforementioned proximal update. In implementation, $(L+1)$-BCD algorithms can be performed in an approximate-proximal

manner—in the first few outer-loop iterations we run the approximate update and then convert to proximal update in the subsequent updates. This approximate-proximal combination exhibits fast convergence and also guarantees stationary-points-achieving convergence as shown previously.

5.5 Numerical Results

In this section, numerical results are presented to verify and compare the performance of the proposed algorithms.

In the following experiments, a wireless sensor network with $L = 3$ sensors is considered. The antenna numbers of the sensors and the fusion center are set as $N_1 = 3, N_2 = 4, N_3 = 5$ and $M = 4$ respectively. All observation matrices K_i are set as identity matrices. The source signal $\mathbf{s}$ has dimension $K = 3$ with zero mean, unit-power and uncorrelated components. The observation noise at each sensor is colored and has covariance matrix $\mathbf{\Sigma}_i = \sigma_i^2 \mathbf{\Sigma}_{0,i}$, $i \in \{1, \cdots, L\}$, where the $J_i \times J_i$ matrix $\mathbf{\Sigma}_{0,i}$ has the Toeplitz structure with its (j,k)-th element $[\mathbf{\Sigma}_{0,i}]_{j,k} = \rho^{|k-j|}$. The parameter ρ is set as 0.5 for all sensors in our test. The transmission power and observation noise at each sensor are set as $P_1 = 2$, $P_2 = 2$, $P_3 = 3$, $\sigma_1^{-2} = 6$dB, $\sigma_2^{-2} = 7$dB and $\sigma_3^{-2} = 8$dB, respectively.

In the test of each algorithm, channel noise level increases from $\mathsf{SNR}_0 = 0$dB to 18dB. For one specific channel noise level, 500 channel realizations $\{\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3\}$ are randomly generated with each matrix entry following standard complex circular Gaussian distribution $\mathcal{CN}(0, 1)$. The mean square error averaged over all 500 random channel

realizations are evaluated as a function of the number of (outer-loop) iterations and the channel SNR.

2-BCD algorithm is implemented by utilizing CVX(with SDPT3 solver) to solve its subproblem (P2). For the essentially cyclic $(L+1)$-BCD algorithm, here we test two special cases: i) the layered BCD algorithm with finite inner-loop iterations, where the inner-loop cyclically updates each beamformer for two times; ii) $(L+1)$B-FG algorithm, where beamformers are cyclically updated with each $\mathbf{F}_i$'s update followed by the calibration of $\mathbf{G}$. That means the variables are updated in an order of $\mathbf{F}_1, \mathbf{G}, \mathbf{F}_2, \mathbf{G}, \cdots$. In one outer-loop it updates each $\mathbf{F}_i$ once and $\mathbf{G}$ for L times. The performance of these two cases are presented in Figure 5.2 and 5.3 respectively. The 2-BCD algorithm is plotted in each figure to serve as a benchmark. On average, the layered-BCD algorithm with finite inner-loop iteration and the $(L+1)$B-FG algorithm converges in 30-40 outer-loop iterations to the identical MSE as that of the 2-BCD algorithm.

The approximate and proximal version (with $\kappa = 1$) of $(L+1)$B-FG algorithm are also tested and presented in Figure 5.4 and 5.5 respectively. As shown in Figure 5.4, the performance of approximate method is surprisingly fast and exhibits excellent performance within only 3 to 5 outer-loop iterations. Comparatively the proximal method, although whose convergence to stationary points can be proved, exhibits a much slower convergence than other algorithms, as shown in Figure 5.5.

In Figure 5.6 the approximate-proximal version of $(L+1)$B-FG is tested. Here in the first 10 outer-loop iterations, approximate version of $(L+1)$B-FG is performed and after that the proximal method is used. As shown in the figure, this combination scheme inherits the fast convergence rate of approximate method and, as proved previously, guarantees convergence to stationary points.

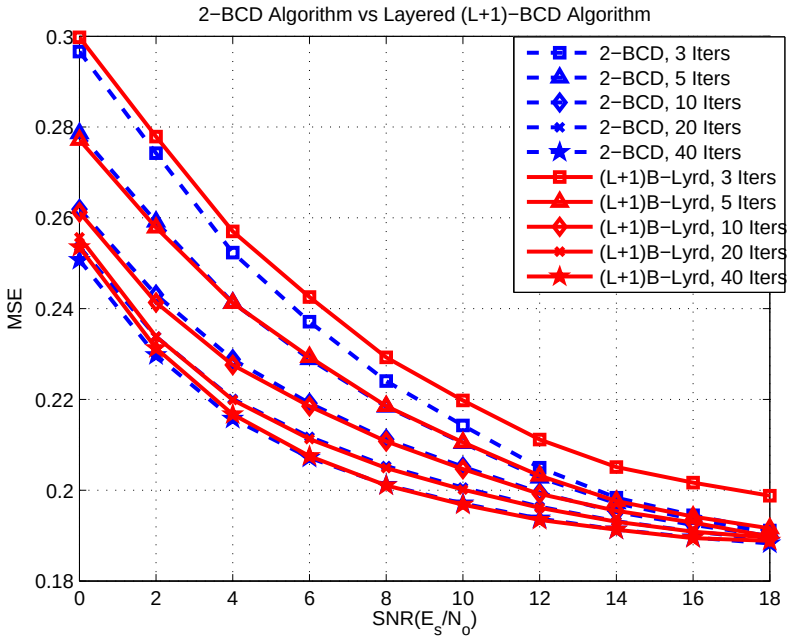

Figure 5.2: MSE Performance of 2-BCD v.s. Layered $(L+1)$-BCD (with 2 inner-loop iterations) Algorithms

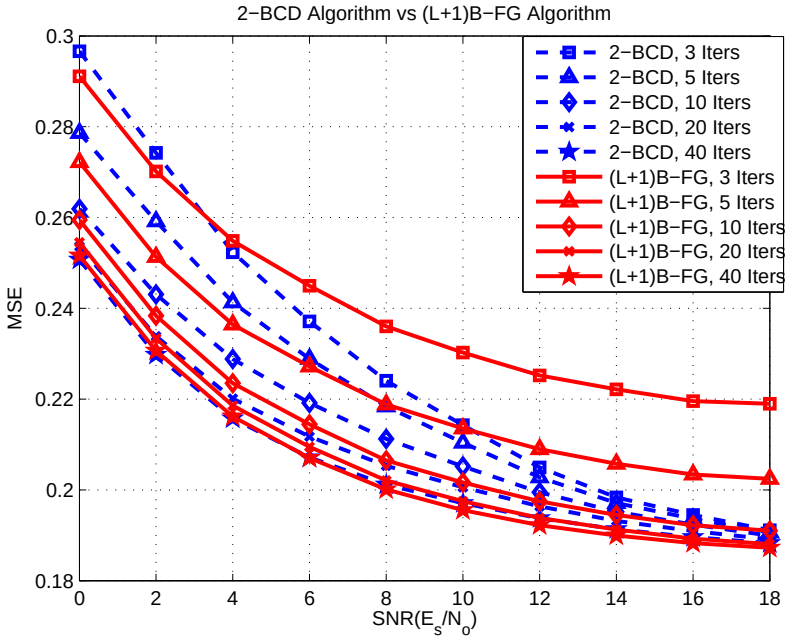

Figure 5.3: MSE Performance of 2-BCD v.s.$(L+1)$B-FG Algorithms

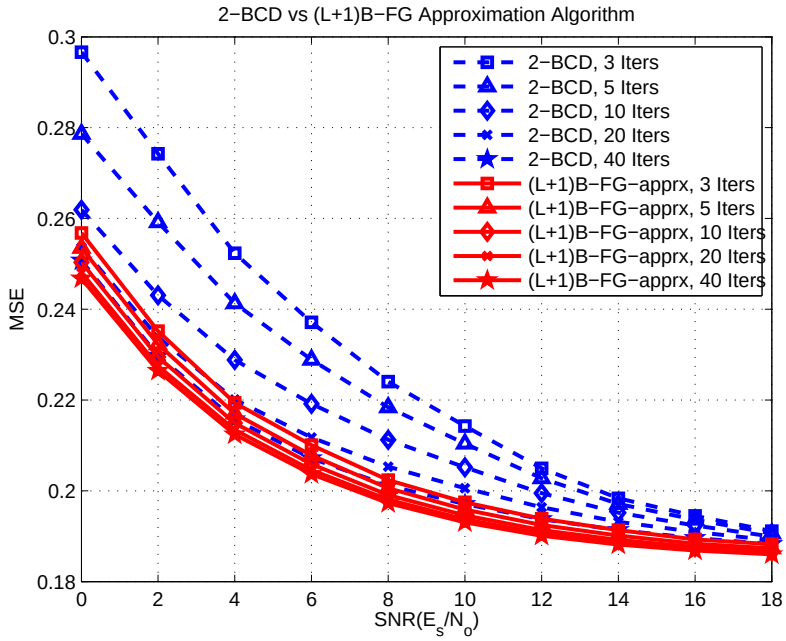

Figure 5.4: MSE Performance of 2-BCD v.s. Approximate $(L+1)$B-FG

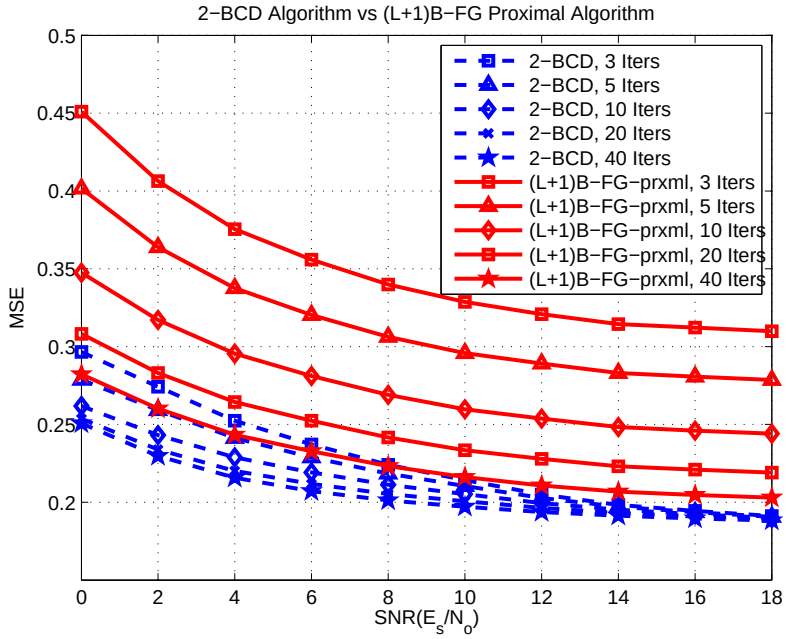

Figure 5.5: MSE Performance of 2-BCD v.s. Proximal $(L+1)$B-FG

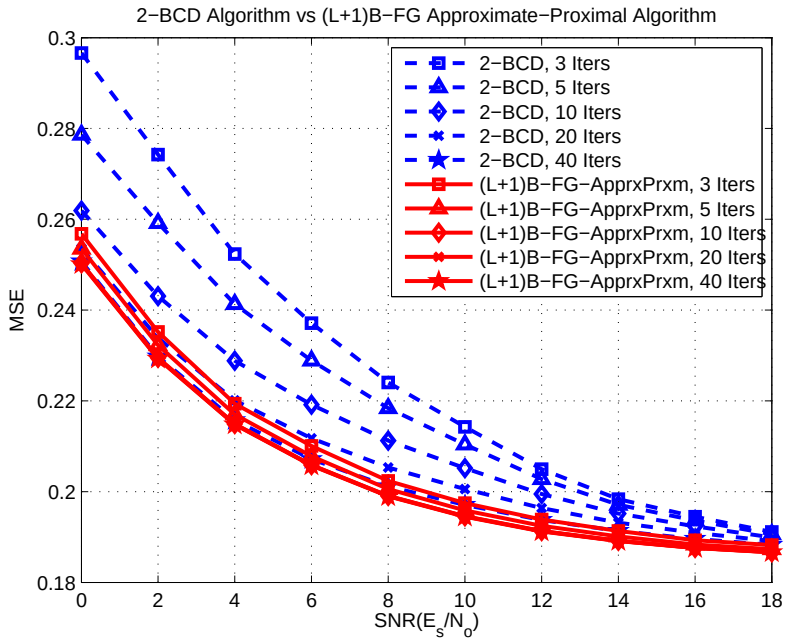

Figure 5.6: MSE Performance of 2-BCD v.s. Approximate-Proximal $(L+1)$B-FG

Next, we take a close look at the convergence behaviors of these algorithms. We set $\mathsf{SNR}_0 = 2$dB and fix the channel by a randomly-generated realization. We randomly generate 10 feasible initial points. We run 2-BCD, $(L+1)$B-FG, proximal $(L+1)$B-FG and approximate-proximal $(L+1)$B-FG algorithms from these 10 random initial points and represent the resultant MSE itineraries in Figures 5.7-5.9. These plots clearly demonstrate that these algorithms are insensitive to initial points and exhibit rather stable converged MSE from different startings. As shown in the figures, different algorithms with random initials finally converge to identical MSE value with different convergence rates. Proximal method has an obviously slower convergence and the approximate-proximal method exhibits fast convergence in the first 3 outer-loop iterations, which coincides with the observations presented in previous figures.

We present in Table 5.1 the average MATLAB running time for different algorithms

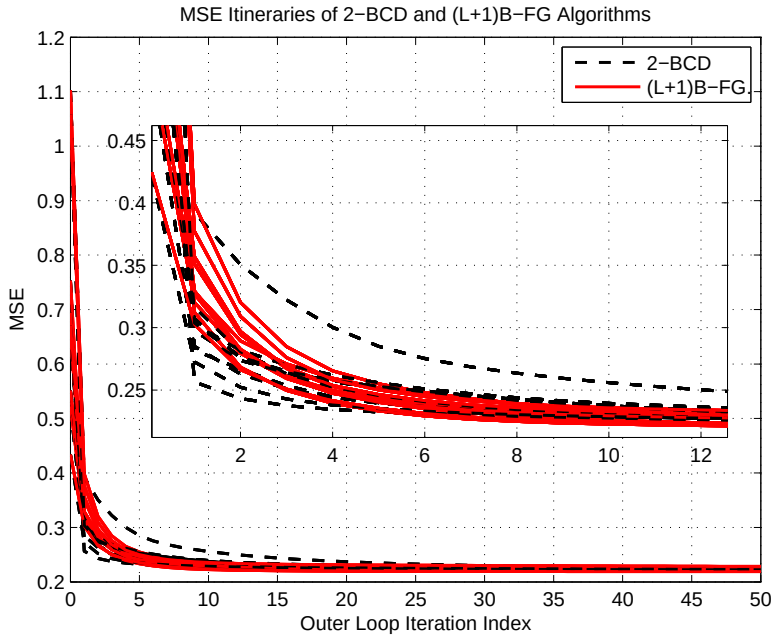

Figure 5.7: MSE Itineraries of 2-BCD v.s. $(L+1)$B-FG Algorithm

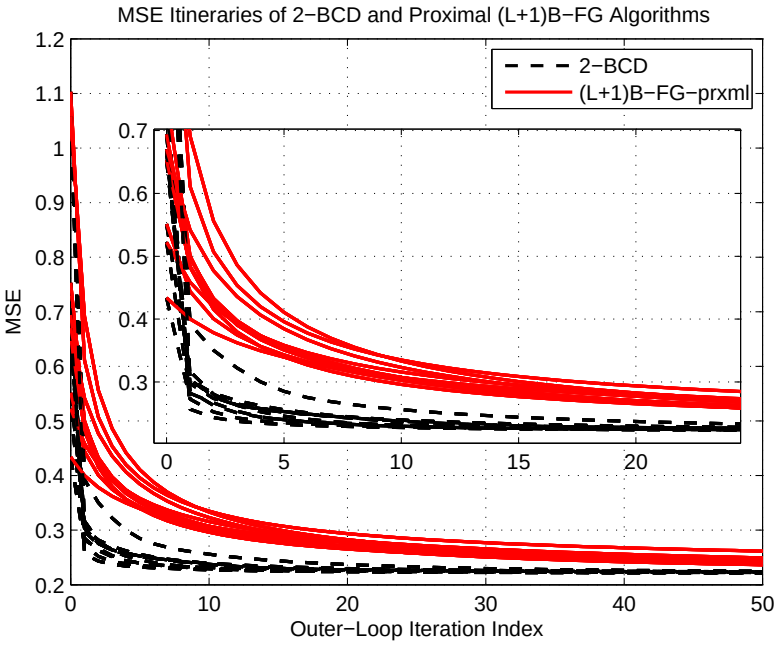

Figure 5.8: MSE Itineraries of 2-BCD v.s. Proximal $(L+1)$B-FG Algorithm

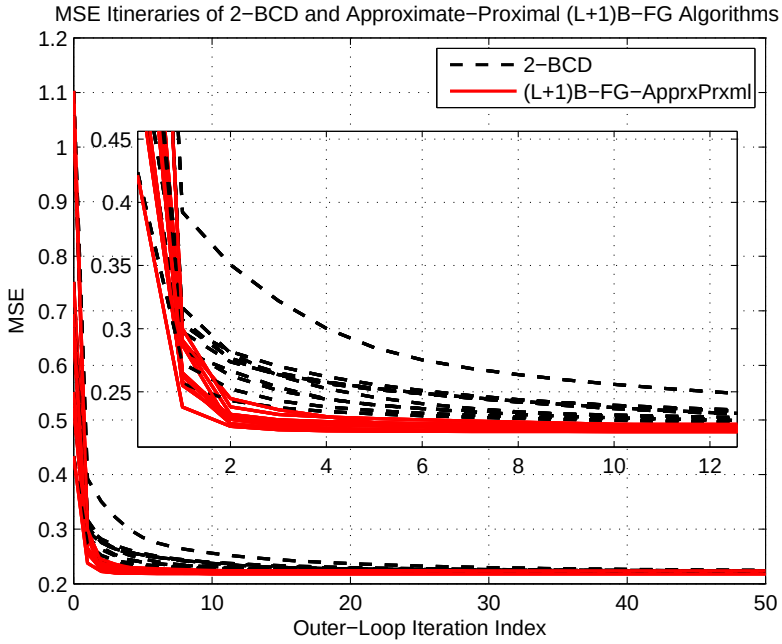

Figure 5.9: MSE Itineraries of 2-BCD v.s. Approximate-Proximal $(L+1)$B-FG Algorithm

(running in a regular laptop). For simplicity, we focus on homogeneous sensor network, where each sensor has the same number of antennae and $J_i = N_i$. Different values of K (size of the source vector) and L (number of sensors) are tested to take into account different problem sizes. The algorithms are run multiple times and the average MATLAB running time per outer-loop iteration is recorded. For the 2-BCD algorithm, CVX is utilized to solve its subproblem and the solver SDPT3 is chosen. In Table 5.1, the average running time of 2-BCD, $(L+1)$B-FG and layered $(L+1)$-BCD algorithm is presented. Note that the approximate, proximal and approximate-proximal $(L+1)$B-FG algorithms have the same complexity as that of $(L+1)$B-FG algorithm. The analytic solutions obtained in Theorem 5.4.1 entitles the essentially cyclic $(L+1)$-BCD algorithm and its variants high efficiency for implementation. However it should be pointed out that the 2-BCD method still has great significance in decentralized optimization for our

Table 5.1: MATLAB Running Time Per (Outer-Loop) Iteration

Dim. \ L	Algorithms	$L=2$	$L=4$	$L=6$	$L=8$
$K=1$	2-BCD	0.2167s	0.2490	0.2987s	0.3500s
$M=3$	$(L+1)$B-FG	0.0026s	0.0066s	0.0120s	0.0189s
$N_i=3$	Lay. BCD	0.0031s	0.0094s	0.0181s	0.0301s
$K=3$	2-BCD	0.2432s	0.3068s	0.3636s	0.4285s
$M=3$	$(L+1)$B-FG	0.0056s	0.0159s	0.0328s	0.0560s
$N_i=3$	Lay. BCD	0.0087s	0.0241s	0.0493s	0.0839s
$K=6$	2-BCD	0.2529s	0.3786s	0.5861s	0.7526s
$M=6$	$(L+1)$B-FG	0.0075s	0.0203s	0.0397s	0.0664s
$N_i=6$	Lay. BCD	0.0116s	0.0319s	0.0622s	0.1031s
$K=9$	2-BCD	0.4352s	0.7956s	1.1401s	1.9593s
$M=9$	$(L+1)$B-FG	0.0120s	0.0302s	0.0557s	0.0902s
$N_i=9$	Lay. BCD	0.0205s	0.0514s	0.0928s	0.1467s

Notes: (i) layered-BCD is run with 2 inner-loop iterations.
(ii) SDPT3 solver of CVX is chosen to implement 2-BCD.

problem. As proved in Theorem 5.3.1, its subproblem (P2) is convex. In fact, by taking advantage of this key property and utilizing multiplier method, the problem (P0) can be solved under the 2-BCD framework, where (P2) is solved in a highly distributed manner with each sensor updating its own beamformer.

5.6 Conclusion

In this chapter we study the joint transceiver design problem for the wireless sensor network under the MSE criterion. Due to the nonconvexity of the original problem, block coordinate descent methods are adopted. A two-block coordinate descent method is first proposed, which decomposes the original problem into two subproblems and alternatively optimizes the linear postcoder and the linear precoders jointly. This 2-BCD algorithm guarantees convergence (of its solution limit points) to stationary points.

We also completely solve the one single beamformer's optimization problem with one power constraint. This conclusion gives birth to highly efficient multiple block coordinate descent methods. We prove the fact that updating the separate beamformer or the linear receiver in any essentially cyclic rule with proximal method can guarantee the convergence to stationary points. Moreover combining approximation with the proximal method significantly improves the convergence rate while maintaining its strong convergence and high efficiency. Extensive numerical results are provided to verify our findings.

5.7 Appendix

5.7.1 Proof of Theorem 5.4.1

Proof. The assumption implies $\left(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H+\mathbf{\Sigma}_i\right) \succ 0$. Therefore $\mathbf{E}_i = \left(\mathbf{K}_i\mathbf{\Sigma_s}\mathbf{K}_i^H+\mathbf{\Sigma}_i\right)\otimes \mathbf{I}_{N_i} \succ 0$. We introduce the following notations

$$\widetilde{\mathbf{f}} \triangleq \mathbf{E}_i^{\frac{1}{2}}\mathbf{f}; \tag{5.26a}$$

$$\mathbf{M}_i \triangleq \mathbf{E}_i^{-\frac{1}{2}}\left(\mathbf{A}_{ii}+\mathbf{C}_i\right)\mathbf{E}_i^{-\frac{1}{2}} = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^H; \tag{5.26b}$$

$$\mathbf{b}_i \triangleq \mathbf{E}_i^{-\frac{1}{2}}\left(\mathbf{B}_i^H\mathbf{g}-\mathbf{q}_i\right); \tag{5.26c}$$

$$\mathbf{p}_i \triangleq \mathbf{U}_i^H\mathbf{b}_i; \tag{5.26d}$$

where the j-th column $\mathbf{u}_{i,j}$ of $\mathbf{U}_i$ is the eigenvector associated with the eigenvalue $\lambda_{i,j} \triangleq [\mathbf{\Lambda}_i]_{j,j}$. Without loss of generality, we assume that the eigenvalues of $\mathbf{M}_i$ are arranged in a decreasing order and that $\mathbf{M}_i$ has rank r_i, $r_i \leq J_iN_i$. In other words $\lambda_{i,1} \geq \cdots \geq$

$\lambda_{i,r_i} > \lambda_{i,r_i+1} = \cdots = \lambda_{i,J_iN_i} = 0.$

Then the problem $(\mathsf{P}2_i')$ is rewritten as

$$(\mathsf{P}3_i) : \min_{\widetilde{\mathbf{f}}_i} \widetilde{\mathbf{f}}_i^H \mathbf{M}_i \widetilde{\mathbf{f}}_i - 2\mathsf{Re}\{\mathbf{b}_i^H \widetilde{\mathbf{f}}_i\}, \tag{5.27a}$$

$$s.t.\ \|\widetilde{\mathbf{f}}_i\|_2^2 \leq P_i. \tag{5.27b}$$

Since $\mathbf{M}_i$ is positive semidefinite, $(\mathsf{P}3_i)$ is convex and is obviously strictly feasible. Thus solving $(\mathsf{P}3_i)$ is equivalent to solving its KKT conditions:

$$\big(\mathbf{M}_i + \mu_i \mathbf{I}\big)\widetilde{\mathbf{f}}_i = \mathbf{b}_i; \tag{5.28a}$$

$$\|\widetilde{\mathbf{f}}_i\|_2^2 \leq P_i; \tag{5.28b}$$

$$\mu_i\big(\|\widetilde{\mathbf{f}}_i\|_2^2 - P_i\big) = 0; \tag{5.28c}$$

$$\mu_i \geq 0. \tag{5.28d}$$

The Lagrangian multiplier μ_i should be either positive or zero, our next discussion focuses on identifying the positivity of μ_i.

Assume that $\mu_i > 0$, then $\big(\mathbf{M}_i + \mu_i \mathbf{I}\big)$ is strictly positive definite and thus invertible. Consequently $\widetilde{\mathbf{f}}_i = \big(\mathbf{M}_i + \mu_i \mathbf{I}\big)^{-1}\mathbf{b}_i$. By the slackness condition (5.28c), the power constraint (5.28b) should be active. Plugging $\widetilde{\mathbf{f}}_i$ into (5.28b) and noting the eigenvalue decomposition in (5.26b), we get

$$\|\widetilde{\mathbf{f}}_i\|^2 = \mathbf{b}_i^H \mathbf{U}_i \big(\mathbf{\Lambda}_i + \mu_i \mathbf{I}\big)^{-2} \mathbf{U}_i^H \mathbf{b}_i = P_i. \tag{5.29}$$

By the definition of $\mathbf{p}_i$ in (5.26d), we rewrite (5.29) as

$$\|\widetilde{\mathbf{f}}_i\|^2 = g_i(\mu_i) = \sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{(\lambda_{i,k}+\mu_i)^2} + \sum_{k=r_i+1}^{J_iN_i} \frac{|p_{i,k}|^2}{\mu_i^2} = P_i. \tag{5.30}$$

Note that here $g_i(\mu_i)$ is a positive, continuous and strictly decreasing function in μ_i.

To identify the positivity of μ_i, the following different cases are considered:

CASE (I)— $\mu_i^\star > 0$ This case further involves two subcases:

case i)— $\exists k \in \{r_i+1, \cdots, J_iN_i\}$ s.t. $|p_{i,k}| \neq 0$:

In this case, it is easily seen that $g_i(\mu_i) \to +\infty$ when $\mu_i \to 0^+$, so $g_i(\mu_i)$ has the range of $\big(0, \infty\big)$ for positive μ_i. So in case i) there always exists a unique positive μ_i satisfying (5.30). Suppose that the unique solution of (5.30) is $\mu_i^\star$. Plugging $\mu_i^\star$ back into the KKT condition (5.28a), we obtain the optimal solution $\widetilde{\mathbf{f}}_i^\star$ as

$$\widetilde{\mathbf{f}}_i^\star = \big(\mathbf{M}_i + \mu_i^\star \mathbf{I}\big)^{-1} \mathbf{b}_i. \tag{5.31}$$

Plugging (5.26) into the above, (5.17) is obtained. It is easily verified that the $\mu_i^\star$ and $\mathbf{f}_i^\star$ in (5.17) satisfy all the KKT conditions in (5.28) and therefore is the optimal solution to ($\mathsf{P2}_i'$).

case ii) $\sum_{k=r_i+1}^{J_iN_i} |p_{i,k}|^2 = 0$ and $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} > P_i$:

In this case, the second part in the summation of $g_i(\mu_i)$ in (5.30) vanishes and $g_i(\mu_i)$ has the bounded range $\left(0, \sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2}\right]$, with its maximum value achieved at $\mu_i = 0$. When $\sum_{k=1}^{r_i} |p_{i,k}|^2 \lambda_{i,k}^2 > P_i$, a positive $\mu_i^\star$ satisfying (5.30) still exists and is unique. Consequently, the optimal solution $\mathbf{f}_i^\star$ can be determined by (5.31) as in the subcase i).

CASE (II)— $\sum_{k=r_i+1}^{J_iN_i} |p_{i,k}|^2 = 0$ and $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} \leq P_i$

In this case, a positive μ_i satisfying KKT conditions does not exist any more, and $\mu_i^\star = 0$. As such, the optimal solution $\mathbf{f}_i^\star$ should satisfy (5.28a):

$$\mathbf{M}_i \widetilde{\mathbf{f}}_i = \mathbf{b}_i. \tag{5.32}$$

We now claim that the above equation (5.32) has a feasible solution. Indeed, this equation is solvable if and only if the right hand side $\mathbf{b}_i$ belongs to the column space $\mathcal{R}\big(\mathbf{M}_i\big)$. Recall that $\mathbf{M}_i$ is Hermitian and has rank r_i; so $\mathcal{R}\big(\mathbf{M}_i\big) = \mathsf{span}\big(\mathbf{u}_{i,1}, \cdots, \mathbf{u}_{i,r_i}\big)$ and the null space of $\mathbf{M}_i$ satisfies $\mathcal{N}\big(\mathbf{M}_i\big) = \mathcal{R}^\perp\big(\mathbf{M}_i\big) = \mathsf{span}\big(\mathbf{u}_{i,r_i+1}, \cdots, \mathbf{u}_{i,J_iN_i}\big)$. In fact, $\mathbb{C}^{J_iN_i} = \mathcal{R}\big(\mathbf{M}_i\big) \oplus \mathcal{N}\big(\mathbf{M}_i\big)$. Invoking the assumption of CASE (II) that $|p_{i,k}| = 0$, $\forall k \in \{r_i + 1, \cdots, J_iN_i\}$ and the definition of $\mathbf{p}_i$, we obtain $p_{i,k} = \mathbf{u}_{i,k}^H \mathbf{b}_i$, $\forall k \in \{r_i + 1, \cdots, J_iN_i\}$. Actually this implies $\mathbf{b}_i \in \mathcal{N}^\perp\big(\mathbf{M}_i\big) = \mathcal{R}\big(\mathbf{M}_i\big)$ and thus the consistency (i.e. the feasibility) of (5.32) is guaranteed.

Next we proceed to analytically identify one special feasible solution of (5.32). Eigenvalue decomposing $\mathbf{M}_i$, (5.32) can be equivalently written as

$$\mathbf{\Lambda}_i \mathbf{U}_i^H \widetilde{\mathbf{f}}_i = \mathbf{p}_i. \tag{5.33}$$

Let $\bar{\mathbf{\Lambda}}_i$ represent the top-left $r_i \times r_i$ sub-matrix of $\mathbf{\Lambda}_i$, i.e. $\mathbf{\Lambda}_i = \mathsf{diag}\big\{\bar{\mathbf{\Lambda}}_i, \mathbf{O}_{J_iN_i - r_i}\big\}$. Let $\bar{\mathbf{U}}_i$ and $\tilde{\mathbf{U}}_i$ represent the left-most r_i columns and the remaining columns of $\mathbf{U}_i$ respectively, i.e. $\mathbf{U}_i = \big[\bar{\mathbf{U}}_i, \tilde{\mathbf{U}}_i\big]$. We can then simplify (5.33) to

$$\bar{\mathbf{\Lambda}}_i \bar{\mathbf{U}}_i^H \widetilde{\mathbf{f}}_i = \mathbf{p}_i. \tag{5.34}$$

Since the columns of $\mathbf{U}_i$ form a set of orthonormal basis for $\mathbb{C}^{J_iN_i}$, $\widetilde{\mathbf{f}}_i$ can be ex-

pressed via columns of $\mathbf{U}_i$ as $\widetilde{\mathbf{f}}_i = \sum_{k=1}^{J_i N_i} \alpha_{i,k}\mathbf{u}_{i,k}$. Noticing the key fact that $\bar{\mathbf{U}}_i^H \mathbf{u}_{i,k} = \mathbf{0}$, $\forall k \in \{r_i + 1, \cdots, J_i N_i\}$, we know that the values of $\{\alpha_{i,r_i+1}, \cdots, \alpha_{i,J_i N_i}\}$ have no impact on (5.34) and can therefore be safely set to zeros to save energy. As for $\alpha_{i,k}$, $\forall k \in \{1, \cdots, r_i\}$, we substitute $\widetilde{\mathbf{f}}_i = \sum_{k=1}^{r_i} \alpha_{i,k}\mathbf{u}_{i,k}$ into (5.34) and obtain

$$\alpha_{i,k} = \lambda_{i,k}^{-1} p_{i,k}, \quad \forall k \in \{1, \cdots, r_i\}. \tag{5.35}$$

Summarizing the above analysis, the optimal solution $\widetilde{\mathbf{f}}_i^\star$ to $(\mathsf{P}3_i)$ is given by

$$\widetilde{\mathbf{f}}_i^\star = \mathbf{U}_i \mathbf{\Lambda}_i^\dagger \mathbf{U}_i^H \mathbf{b}_i, \tag{5.36}$$

with $\mathbf{\Lambda}_i^\dagger$ being the Moore-Penrose pseudoinverse of $\mathbf{\Lambda}_i$ given as $\mathsf{diag}\big\{\bar{\mathbf{\Lambda}}_i^{-1}, \mathbf{O}_{J_i N_i - r_i}\big\}$. Matrix theory suggests that an arbitrary matrix $\mathbf{X}$ with its singular value decomposition (SVD) given by $\mathbf{X} = \mathbf{U}_\mathbf{X} \mathbf{\Lambda}_\mathbf{X} \mathbf{V}_\mathbf{X}^H$ has its unique Moore-Penrose pseudoinverse $\mathbf{X}^\dagger = \mathbf{V}_\mathbf{X} \mathbf{\Lambda}_\mathbf{X}^\dagger \mathbf{U}_\mathbf{X}^H$, where $\mathbf{U}_\mathbf{X}$ and $\mathbf{V}_\mathbf{X}$ are left and right singular square matrices, respectively, and $\mathbf{\Lambda}_\mathbf{X}$ is a diagonal matrix with appropriate dimensions. Hence, (5.36) can be equivalently written as

$$\widetilde{\mathbf{f}}_i^\star = \mathbf{M}_i^\dagger \mathbf{b}_i. \tag{5.37}$$

Obviously $\mu_i^\star = 0$, and $\mu_i^\star$ and $\widetilde{\mathbf{f}}_i^\star$ satisfy the KKT conditions (5.28a), (5.28c) and (5.28d). What remains to be shown is that $\widetilde{\mathbf{f}}_i^\star$ satisfies the power constraint. We verify this using (5.35) and get

$$\|\widetilde{\mathbf{f}}_i^\star\|^2 = \sum_{k=1}^{r_i} |\alpha_{i,k}|^2 = \sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} \leq P_i, \tag{5.38}$$

where the inequality in the above follows the assumption of CASE (II). Plugging (5.26) into (5.37), (5.18) is obtained. The proof is complete. □

5.7.2 Proof of Theorem 5.4.3

Proof. This proof is inspired by Proposition 2.7.1 in [63]. To simplify the following exposition, we define $\mathbf{x} \triangleq [\mathbf{x}_1^T, \cdots, \mathbf{x}_{L+1}^T] = [\mathbf{f}_1^T, \cdots, \mathbf{f}_L^T, \mathbf{g}^T]$ and $\mathbf{x} \in \mathcal{X} \triangleq \mathcal{X}_1 \times \cdots \times \mathcal{X}_{L+1}$ with $\mathcal{X}_i = \left\{\mathbf{f}_i \in \mathbb{C}^{J_i N_i} \middle| \mathbf{f}_i^H \mathbf{E}_i \mathbf{f}_i \leq P_i\right\}$, for $i = 1, \cdots, L$ and $\mathcal{X}_{L+1} = \mathbb{C}^{KM}$. For any specific essentially cyclic update BCD algorithm, we assume that it starts from an initial feasible solution $\mathbf{x}^{(0)} \triangleq [\mathbf{x}_1^{T(0)}, \cdots, \mathbf{x}_{L+1}^T{}^{(0)}]$ and the iteration index (k) increases by one after any block's update. Denote $\mathbf{x}_i^{(k)}$ as the i-th block of $\mathbf{x}^{(k)}$ and $\mathbf{x}_{\bar{i}} = [\mathbf{x}_1, \cdots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \cdots, \mathbf{x}_{L+1}]$, $i \in \{1, \cdots, L+1\}$, $i \in \{1, \cdots, L+1\}$. Assume that T is a period of the essentially cyclic update rule and $\{t_1, \cdots, t_T\}$, with $t_j \in \{1, \cdots, L+1\}$ $\forall j \in \{1, \cdots, T\}$, as the indices of the updated blocks in a period in order. If $\mathbf{x}_{t_j}$ is updated in the (k)-th iteration, then $\mathbf{x}_{t_{j\oplus 1}}$ is updated in the $(k+1)$-th iteration. Define $j \oplus 1 \triangleq j (\mathrm{mod}\ T) + 1$, $\forall j \in \{1, \cdots, T\}$ and $j \oplus m$ as $j \oplus 1$ by m times.

By repeatedly invoking Bolzano-Weierstrass theorem to $\mathbf{f}_i$ to $\mathbf{f}_L$ and noticing that $\mathbf{g}$ is updated in closed form by equation (5.9), the existence of limit points of $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ can be proved.

Then we prove that $\mathsf{MSE}^{(k)}$ is decreasing. If $\mathbf{x}_{L+1}$(or $\mathbf{g}$) is updated in the $(k+1)$-th iteration, then (P1) is solved and thus MSE is decreasing. Assume that in the $(k+1)$-th iteration, the $(t_{j\oplus 1})$-th block is updated, $t_{j\oplus 1} \in \{1, \cdots, L\}$. Then

$$\mathbf{x}_{t_{j\oplus 1}}^{(k+1)} = \underset{\mathbf{x}_{t_{j\oplus 1}} \in \mathcal{X}_{t_{j\oplus 1}}}{\arg\min} . \ \mathsf{MSE}\big(\mathbf{x}_{t_{j\oplus 1}} \big| \mathbf{x}_{\overline{t_{j\oplus 1}}}^{(k)}\big) + \kappa \big\| \mathbf{x}_{t_{j\oplus 1}} - \mathbf{x}_{t_{j\oplus 1}}^{(k)} \big\|_2^2.$$

Since $\mathbf{x}_{t_j}^{(k)}$ is feasible, it should give no smaller objective than $\mathbf{x}_{t_j}^{(k+1)}$ for the above problem. This implies

$$\mathsf{MSE}\big(\mathbf{x}^{(k+1)}\big) \leq \mathsf{MSE}\big(\mathbf{x}^{(k)}\big) - \kappa \big\|\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}\big\|_2^2 \leq \mathsf{MSE}\big(\mathbf{x}^{(k)}\big).$$

Thus $\mathsf{MSE}^{(k)}$ is decreasing. At the same time notice that MSE should be nonnegative, thus $\mathsf{MSE}^{(k)}$ converges.

Next we prove that any limit point is stationary. Assume that a subsequence of solution $\mathbf{x}^{(k_j)}$ converges to a limit point $\bar{\mathbf{x}} \triangleq [\bar{\mathbf{x}}_1^T, \cdots, \bar{\mathbf{x}}_{L+1}^T]$. Since there are finite blocks, we assume the block $i \in \{1, \cdots, L+1\}$ is updated infinitely many times and assume that $i = t_l$ for some $l \in \{1, \cdots, T\}$. It should be noted that such l may be non-unique and arbitrary one can be chosen to do the job.

We assert that $\mathbf{x}^{(k_j+1)} \to \bar{\mathbf{x}}$, i.e. $\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} \to \bar{\mathbf{x}}_{t_{l\oplus 1}}$. This claim can be proved in two cases—i) $t_{l\oplus 1} = L+1$ and ii) $t_{l\oplus 1} \in \{1, \cdots, L\}$.

i) $t_{l\oplus 1} = L+1$. Notice that $\mathbf{x}_{L+1} = \mathbf{g}$ is updated in a closed form (5.9), which is a continuous function of $[\mathbf{x}_1^T, \cdots, \mathbf{x}_L^T]$. Since $\mathbf{x}_{\overline{L+1}}^{(k_j)}$ converges, by taking $j \to \infty$, $\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)}$ should converge to some limit, i.e. $\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} \to \widetilde{\mathbf{x}}_{L+1}$. Notice that $\mathsf{MSE}^{(k)}$ converges, so $\mathsf{MSE}\big(\bar{\mathbf{x}}_{\overline{L+1}}, \bar{\mathbf{x}}_{L+1}\big) = \mathsf{MSE}\big(\bar{\mathbf{x}}_{\overline{L+1}}, \widetilde{\mathbf{x}}_{L+1}\big)$. This means both $\bar{\mathbf{x}}_{L+1}$ and $\widetilde{\mathbf{x}}_{L+1}$ are solutions to the problem (P1) with sensors' beamformers $[\bar{\mathbf{x}}_1^T, \cdots, \bar{\mathbf{x}}_L^T]$ given. Since (P1) is strictly convex and thus has unique solution, we conclude $\widetilde{\mathbf{x}}_{L+1} = \bar{\mathbf{x}}_{L+1}$. So $\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} \to \bar{\mathbf{x}}_{t_{l\oplus 1}}$ holds for the case $t_{l\oplus 1} = L+1$.

ii) $t_{l\oplus 1} \in \{1, \cdots, L\}$. By contradiction, we assume that $\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)}$ does not converge to $\bar{\mathbf{x}}_{t_{l\oplus 1}}$. By denoting $\gamma^{(k_j)} \triangleq \|\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} - \bar{\mathbf{x}}_{t_{l\oplus 1}}\|_2$ and possibly restricting to a subsequence, we assume that there exists a $\bar{\gamma} > 0$ such that $\gamma^{(k_j)} \geq \bar{\gamma}$ for all j. Let $\mathbf{s}_l^{(k_j)} = (\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} -$

$\mathbf{x}_{t_{l\oplus 1}}^{(k_j)})/\gamma^{(k_j)}$. Since $\mathbf{s}^{(k_j)}$ is bounded, by Bolzano-Weierstrass theorem and restricting to a subsequence, we assume that $\mathbf{s}^{(k_j)} \to \bar{\mathbf{s}}$. Then we obtain

$$\begin{aligned}
&\mathsf{MSE}\big(\mathbf{x}^{(k_j+1)}\big) = \mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} \big| \mathbf{x}_{\overline{t_{l\oplus 1}}}^{(k_j)}\big) && (5.39)\\
\leq &\mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} \big| \mathbf{x}_{\overline{t_{l\oplus 1}}}^{(k_j)}\big) + \kappa \big\| \mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} - \mathbf{x}_{t_{l\oplus 1}}^{(k_j)} \big\|_2^2 && (5.40)\\
= &\mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j)} + \gamma^{(k_j)}\mathbf{s}^{(k_j)} \big| \mathbf{x}_{\overline{t_{l\oplus 1}}}^{(k_j)}\big) + \kappa \big\| \gamma^{(k_j)}\mathbf{s}^{(k_j)} \big\|_2^2 && (5.41)\\
\leq &\mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j)} + \epsilon\bar{\gamma}\mathbf{s}^{(k_j)} \big| \mathbf{x}_{\overline{t_{l\oplus 1}}}^{(k_j)}\big) + \kappa \big\| \epsilon\bar{\gamma}\mathbf{s}^{(k_j)} \big\|_2^2, \forall \epsilon \in [0,1] && (5.42)\\
\leq &\mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j)} \big| \mathbf{x}_{\overline{t_{l\oplus 1}}}^{(k_j)}\big) = \mathsf{MSE}\big(\mathbf{x}^{(k_j)}\big), && (5.43)
\end{aligned}$$

where the last two inequalities follow the fact that $\mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}} \big| \mathbf{x}_{\overline{t_{l\oplus 1}}}^{(k_j)}\big) + \kappa \|\mathbf{x}_{t_{l\oplus 1}} - \mathbf{x}_{t_{l\oplus 1}}^{(k_j)}\|_2^2$ is strictly convex and attains the minimum at point $\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)}$. Noting $\mathsf{MSE}^{(k_j)}$ converges and letting $j \to \infty$, we obtain

$$\begin{aligned}
\mathsf{MSE}\big(\bar{\mathbf{x}}\big) &\leq \mathsf{MSE}\big(\bar{\mathbf{x}}_{t_{l\oplus 1}} + \epsilon\bar{\gamma}\bar{\mathbf{s}} \big| \bar{\mathbf{x}}_{\overline{t_{l\oplus 1}}}\big) + \kappa\epsilon^2\bar{\gamma}^2 \\
&\leq \mathsf{MSE}\big(\bar{\mathbf{x}}\big), \quad \forall \epsilon \in [0,1], && (5.44)
\end{aligned}$$

which immediately implies

$$\mathsf{MSE}\big(\bar{\mathbf{x}}_{t_{l\oplus 1}} + \epsilon\bar{\gamma}\bar{\mathbf{s}} \big| \bar{\mathbf{x}}_{\overline{t_{l\oplus 1}}}\big) + \kappa\epsilon^2\bar{\gamma}^2 = \mathsf{MSE}\big(\bar{\mathbf{x}}\big), \forall \epsilon \in [0,1]. \quad (5.45)$$

However the above is impossible. Notice that $\mathsf{MSE}\big(\bar{\mathbf{x}}_{t_{l\oplus 1}} + \epsilon\bar{\gamma}\bar{\mathbf{s}} \big| \bar{\mathbf{x}}_{\overline{t_{l\oplus 1}}}\big)$ is a quadratic function of ϵ with nonnegative quadratic coefficient and $\bar{\gamma}, \kappa > 0$. Thus the left hand side(LHS) of equation (5.45) is a strictly convex quadratic function of ϵ, which has at most two different ϵ giving the function value of $\mathsf{MSE}(\bar{\mathbf{x}})$. Contradiction has been reached.

In the above we have proved that $\mathbf{x}^{(k_j+1)} \to \bar{\mathbf{x}}$. Next we show that $\nabla_{\mathbf{x}_{t_{l\oplus 1}}} \mathsf{MSE}\big(\bar{\mathbf{x}}\big)^T \big(\mathbf{x}_{t_{l\oplus 1}} - \bar{\mathbf{x}}_{t_{l\oplus 1}}\big) \geq 0, \forall \mathbf{x}_{t_{l\oplus 1}} \in \mathcal{X}_{t_{l\oplus 1}}$, which is also proved in two cases:

When $t_{l\oplus 1} \in \{1, \cdots, L\}$, we have

$$\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} = \underset{\mathbf{x}_{t_{l\oplus 1}} \in \mathcal{X}_{t_{l\oplus 1}}}{\arg\min} . \mathsf{MSE}\Big(\mathbf{x}_{t_{l\oplus 1}} \big| \mathbf{x}_{t_{l\oplus 1}}^{(k_j)}\Big) + \kappa \big\| \mathbf{x}_{t_{l\oplus 1}} - \mathbf{x}_{t_{l\oplus 1}}^{(k_j)} \big\|_2^2.$$

By optimality condition, the above implies

$$\begin{aligned} &\nabla_{\mathbf{x}_{t_{l\oplus 1}}} \mathsf{MSE}\big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} \big| \mathbf{x}_{t_{l\oplus 1}}^{(k_j)}\big)^T \big(\mathbf{x}_{t_{l\oplus 1}} - \mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)}\big), \\ &+ 2\kappa \big(\mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)} - \mathbf{x}_{t_{l\oplus 1}}^{(k_j)}\big)^T \big(\mathbf{x}_{t_{l\oplus 1}} - \mathbf{x}_{t_{l\oplus 1}}^{(k_j+1)}\big) \geq 0, \forall \mathbf{x}_{t_{l\oplus 1}} \in \mathcal{X}_{t_{l\oplus 1}}. \end{aligned} \tag{5.46}$$

Let $j \to \infty$ in the above equation and note that MSE is continuously differentiable, we obtain

$$\nabla_{\mathbf{x}_{t_{l\oplus 1}}} \mathsf{MSE}\big(\bar{\mathbf{x}}\big)^T \big(\mathbf{x}_{t_{l\oplus 1}} - \bar{\mathbf{x}}_{t_{l\oplus 1}}\big) \geq 0, \forall \mathbf{x}_{t_{l\oplus 1}} \in \mathcal{X}_{t_{l\oplus 1}}. \tag{5.47}$$

When $t_{l\oplus 1} = L + 1$, the above reasoning still works except that the proximal term is absent (i.e. $\kappa = 0$). So we also obtain $\nabla_{\mathbf{x}_{t_{l\oplus 1}}} \mathsf{MSE}\big(\bar{\mathbf{x}}\big)^T \big(\mathbf{x}_{t_{l\oplus 1}} - \bar{\mathbf{x}}_{t_{l\oplus 1}}\big) \geq 0$.

Now replace the subsequence $\{k_j\}$ with $\{k_j + 1\}$, $t_{l\oplus 1}$ with $t_{l\oplus 2}$ and utilize the verbatim argument as above, we can prove

$$\nabla_{\mathbf{x}_{t_{l\oplus 2}}} \mathsf{MSE}\big(\bar{\mathbf{x}}\big)^T \big(\mathbf{x}_{t_{l\oplus 2}} - \bar{\mathbf{x}}_{t_{l\oplus 2}}\big) \geq 0, \forall \mathbf{x}_{t_{l\oplus 2}} \in \mathcal{X}_{t_{l\oplus 2}}. \tag{5.48}$$

Repeating this argument for $(T-1)$ times and recalling that for essentially cyclic update

rule, $\{t_{l\oplus 1},\cdots,t_{l\oplus T}\}=\{1,\cdots,L\}$, we have proved that

$$\nabla_{\mathbf{x}_i}\mathsf{MSE}\big(\bar{\mathbf{x}}\big)^T\big(\mathbf{x}_i-\bar{\mathbf{x}}_i\big)\geq 0,\forall\mathbf{x}_i\in\mathcal{X}_i,\forall i\in\{1,\cdots,L+1\}. \tag{5.49}$$

Summing up the above $(L+1)$ inequalities, we obtain

$$\nabla_{\mathbf{x}}\mathsf{MSE}\big(\bar{\mathbf{x}}\big)^T\big(\mathbf{x}-\bar{\mathbf{x}}\big)\geq 0,\forall\mathbf{x}\in\mathcal{X}. \tag{5.50}$$

So $\bar{\mathbf{x}}$ is actually a stationary point of (P0). □

Chapter 6

Joint Transceiver Design towards MI Maximization for Wireless Sensor Network

6.1 Introduction

In last chapter, we have discussed joint transceiver design problem in centralized wireless sensor networks, where mean square error(MSE) is adopted as our system's performance metric. As we have discussed before, MSE is a standard metric for signal estimation problem, which measures the average deviation of the estimator output compared to the original source. For system's performance evaluation, we usually have many different perspectives. Now the wireless sensor network is a communication system in the first place. It transfers the information of the source(the observed events

here) to the fusion center(FC). So the information rate that we can transfer from the source to FC is of great concerns. How to jointly design the transceivers to maximize the information of the source revealed at the FC is an interesting and important problem. This is actually the motivation of this chapter. Here we research the problem of jointly designing transceivers to maximize the mutual information of the source at the FC.

The contribution of this chapter is that for MI optimization, inspired by the seminal idea of weighted minimum mean square error (WMMSE) method in [60, 73], we introduce weight matrix and a virtual FC receiver as intermediate variables(receiver does not impact MI, so the original problem is independent of FC receiver). Here we develop two BCD algorithms. Firstly, we decompose the MI problem into three subproblems—one subproblem to update the virtual FC receiver, one subproblem to update the weight matrix and the third subproblem to jointly optimize the entire beamformers of all sensors. Both of the two former subproblems have closed form solutions and the third one is convex, which actually can be converted into a second order cone programming(SOCP) problem, and thus efficiently solved by standard numerical solvers. The convergence of this 3-block BCD algorithm is carefully examined and we prove that its limit points are stationary. Secondly, following the route of the first algorithm, we proceed to further decompose its third subproblem into multiple atom problems, with each atom problem optimizing one separate sensor's beamformer. We carefully examine its optimality conditions and obtain the almost closed form solution. It should be noted that, although the technique of checking KKT condition for each separate beamformer is rather standard and has also been adopted in several previous papers (e.g. [58], [61], [53] and [60]), we are able to fully solve the problem by clearly describing the solution structure and deriving the exact closed-form solutions. To be specific, we explicitly obtain the equivalent conditions for judging the positiveness of the Lagrange multipliers, and, in the

case of zero-Lagrange-multipliers, we derive the solution via pseudoinverse. These exact results, and especially the case of the zero-Lagrange-multiplier, are not discussed previously in the literature.

Recently the wireless sensor network (WSN) has attracted great attentions due to its wide applications in practice [48–54, 56, 57, 71, 72]. A typical wireless sensor network has multiple sensors which are spacially distributed and wirelessly connected. Sensors in the same neighborhood monitor the same physical event or measure some common environmental parameters and transmit their (usually contaminated) observations to a preassigned fusion center (FC) to perform further data processing and fusion. The goal of data transmission and fusion in wireless sensor networks can be achieved more effectively by leveraging multiple antenna and linear beamforming techniques. It is always an interesting and meaningful problem to collaboratively design the beamformers so that the wireless sensor network can reliably transmit and recover the observed signals.

Magnitudes of studies have been performed on the beamforming design problem in wireless sensor networks and solutions are provided from various perspectives. For example, the papers [48–51] aim at designing effective beamformers for signal compression. [48] and [49] consider the perfect channel case, i.e. there exists no noise or fading in transmission from the sensors to the FC. Although the perfect channel assumption in [48] and [49] excludes power constraints and greatly simplifies the problem, it is too restrictive for wireless settings. More practical models with noisy channels are considered in [50–54, 56, 57, 71, 72] [50] considers the problem of transmission and fusion of scalar source signal for noisy multiple access channels (MAC), where all sensors share one total transmission power. In practice total power constraint could still be stringent since the sensors are usually sufficiently distributed within a large area and therefore

power sharing is hard to realize. [51] studies noisy and fading channels and separate power constraint for each sensor, under the assumptions that all channel matrices are square and nonsingular. [57] considers the special case of scalar source signal, where separate power constraints and noisy channels are assumed.

Compared to the forementoined literature, the most generic model for centralized wireless sensor network is first introduced in [52]. The model proposed by [52] considers fading and noisy channels, separate power constraint for each sensor and both orthogonal and coherent MAC. Besides the above, no additional assumptions are imposed on the dimension/rank of beamformers or channel matrices, i.e. beamformers can be compressive, redundancy-added or rate-1 and channel matrices can be slim, flat or square (singular or nonsingular). Due to the difficulty of the problem, [52] provides solutions to several important special cases subsumed in the generic model for coherent MAC, including the scalar source signal case, the noiseless sensor-FC channel case and the no-intersymbol-interference (no-ISI) noisy channel case. Following the exact generic model in [52], [53] develops an iterative block coordinate descent (BCD) method that is applicable to any general case for coherent MAC. Recently various strong-convergence-guaranteed BCD-based algorithms have been proposed in [71], which can solve the most generic model in [52] for coherent MAC and subsume the algorithm by [53] as a specialized realization. All of the above mentioned papers [48–53, 57, 71] adopt mean square error (MSE) as performance metric.

Besides the MSE criterion, signal to noise ratio (SNR) is another crucial and commonly used metric for scalar signal recovery. For the coherent MAC wireless sensor networks proposed in [52], joint beamforming design towards maximizing SNR is reported in [56] and [72]. Recently joint beamforming design to maximize mutual information

(MI) for orthogonal MAC is considered in [54].

It worth noting that the beamforming design problems in MIMO multi-sensor decision-fusion system are closely related with those in other multi-agent communication networks, e.g. MIMO multi-relay and multiuser communication systems. Plenty of exciting results exist in literature, see, for example, [58, 59, 61] and the reference therein.

The contribution of this chapter is follows:

1) In this chapter we research the joint beamforming design in coherent MAC wireless sensor network towards MI maximization. Just as the MSE and SNR metric, MI is also a very meaningful design criterion, which is commonly adopted in communication theory to evaluate the average information transmission rate of a system. In wireless sensor network, MI represents the average information of the source signal which can be extracted at the fusion center from the sensors' observations for each use of channel. Compared to the great deal of existing literature focusing on MSE, however, not many results have been reported on MI optimization in the WSN context due to its difficult nature. One recent inspiring paper [54] provides the beamforming solution to maximize MI in the orthogonal MAC wireless sensor network. As will be seen, the original MI optimization problem in coherent MAC wireless sensor networks is also a highly nonconvex hard problem and efficient solutions are meaningful and desirable.

2) Inspired by the seminal idea of weighted minimum mean square error (WMMSE) method in [73] and [60], we introduce a weight matrix and a virtual FC receiver as intermediate variables (the original MI maximization problem does not assume the presence of linear filter at FC, since, according to the data-processing inequality in [77], MI will never increase whatever processing procedure is performed at the receiver) and develop

block coordinate ascent (BCA) algorithms to efficiently solve the original problem. Here we decompose the MI problem into three subproblems—one subproblem to update the virtual FC receiver, one subproblem to update the weight matrix and the third one to jointly optimize the entire beamformers of all sensors. The two former subproblems have closed form solutions and the third one can be proved to be a standard second order cone programming (SOCP) problem. The convergence analysis shows that the limit points of our solutions satisfy Karush-Kuhn-Tucker (KKT) conditions of the original MI maximization problem.

3) Besides the above 3 block BCA algorithm, we also come out a multiple block BCA algorithm, which has closed form solutions (possibly up to a simple bisection search) for each subproblem and is consequently highly efficient for implementation and not reliant on numerical solvers. Moreover we show that, in special circumstance, fully analytical update is even possible for the multiple BCA algorithm. Complexity of this algorithm is examined and extensive numerical results show that this multiple block BCA algorithm exhibits quite good convergence performance.

The rest of the chapter is organized as follows: Section 6.2 introduces the system model of the coherent MAC wireless sensor network and formulates the joint beamforming problem towards maximizing mutual information. In section 6.3 we propose two BCA based algorithms to solve our original problem, with the convexity, closed form solutions and convergence being discussed in full details. Section 6.4 provides numerical experiment results. Section 6.5 concludes the article.

Notations: In the sequel, we use bold lowercase letters to denote complex vectors and bold capital letters to denote complex matrices. $\mathbf{0}$, $\mathbf{O}_{m\times n}$, and $\mathbf{I}_m$ are used to denote zero vectors, zero matrices of dimension $m \times n$, and identity matrices of order

m respectively. $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$, and $\mathbf{A}^\dagger$ are used to denote the transpose, the conjugate, the conjugate transpose(Hermitian transpose), and the Moore-Penrose pseudoinverse respectively of an arbitrary complex matrix $\mathbf{A}$. $\mathsf{Tr}\{\cdot\}$ denotes the trace operation of a square matrix. $|\cdot|$ denotes the modulus of a complex scalar, and $\|\cdot\|_2$ denotes the l_2-norm of a complex vector. $vec(\cdot)$ means vectorization operation of a matrix, which is performed by packing the columns of a matrix into a long one column. $\otimes$ denotes the Kronecker product. $\mathsf{Diag}\{\mathbf{A}_1, \cdots, \mathbf{A}_n\}$ denotes the block diagonal matrix with its i-th diagonal block being the square complex matrix $\mathbf{A}_i$, $i \in \{1, \cdots, n\}$. $\mathcal{H}_+^n$ and $\mathcal{H}_{++}^n$ represent the cones of positive semidefinite and postive definite matrices of dimension n respectively. Here $\succeq 0$ and $\succ 0$ denote that a square complex matrix belongs to $\mathcal{H}_+^n$ and $\mathcal{H}_{++}^n$ respectively. $\mathsf{Re}\{x\}$ means taking the real part of a complex value x.

6.2 System Model

Here we consider the centralized wireless sensor network as illustrated in Fig.6.1.

This system has L sensors and one fusion center. We assume that all sensors and the FC are equipped with multi-antenna. Denote the number of the antennae of the i-th sensor as N_i, $i \in \{1, \cdots, L\}$, and that of the FC as M. The source signal $\mathbf{s}$ is a complex vector of dimension K, i.e. $\mathbf{s} \in \mathbb{C}^{K\times 1}$. Each sensor utilizes a linear beamformer(transmiter/precoder) $\mathbf{F}_i \in \mathbb{C}^{N_i \times K}$ to transmit its observed data. In the beamforming problem to optimize MSE or SNR, a linear receiver(postcoder) is usually employed on the side of fusion center. In fact the presence of linear receiver at the FC leads to joint optimization of transmitters and receivers and can greatly improve the performance metric in terms of MSE or SNR. However linear receiver is not considered

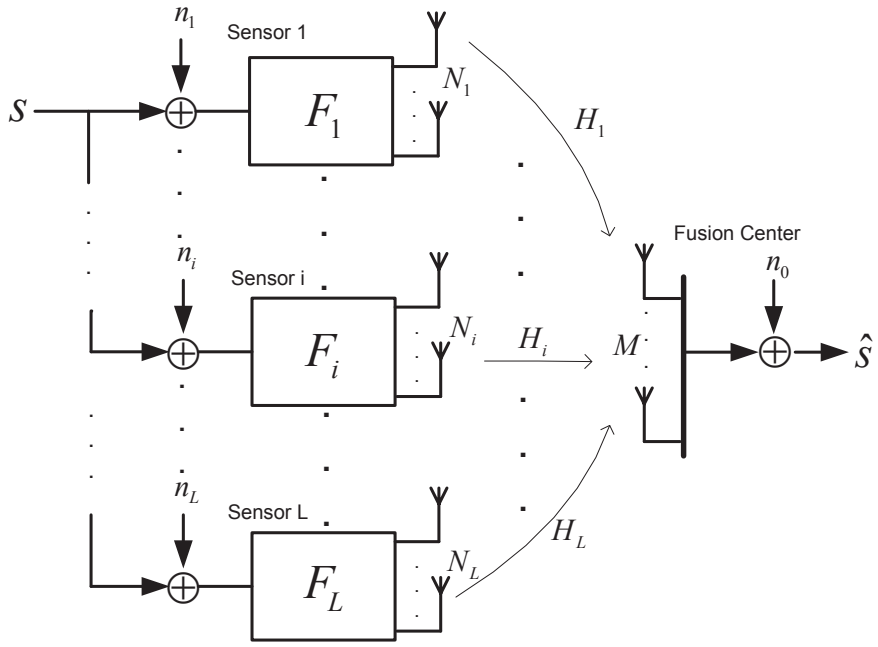

Figure 6.1: Multi-Sensor System Model

for MI maximization problem. According to information nonincreasing principle, any kinds of processing at the receiver will not increase the mutual information between the source and receiver. Thus, without loss of optimality, no filter is necessary at FC.

Here we adopt the assumption that the source signal follows zero mean circularly symmetric complex Gaussian distribution, i.e. $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_s})$ with $\mathbf{\Sigma_s}$ being positive definite. The meaning of the Gaussian signaling assumption has many folds as follows: generally MI lacks of analytical expression and the Gaussian signaling is one of the few exceptions having elegant closed form, which makes our problem analyzable. Moreover, for given source covariance and Gaussian vector channel, Gaussian source maximizes the mutual information [77], i.e. achieves the channel capacity. So in practice sensors can perform signaling transformation to approximate the transformed signals' distribution to Gaussian distribution to improve transmission efficiency [75, 76]. At least,

Gaussian source can provide an upper bound of the reduced uncertainty of the source at the fusion center. At the same time by central limit theory, the Gaussian signaling can serve as a good approximation for a large number of observations which follow independent and identical distribution.

The channel status $\{\mathbf{H}_i\}_{i=1}^{L}$ are assumed to be known at the receiver, which can be achieved by standard channel estimation technique via pilots. We denote $\mathbf{H}_i \in \mathbb{C}^{M\times N_i}$ as the channel coefficients from the i-th sensor to the fusion center. Due to interference from surroundings or thermal noise from the sensor device, the observed signals at the sensors are typically contaminated. We assume that the corruptions are additive zero mean circularly symmetric complex Gaussian noise, i.e. $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_i), i \in \{1, \cdots, L\}$ with $\mathbf{\Sigma}_i \in \mathbb{C}^{K\times K}$ being covariance matrix. Since the sensors are spatially distributed, it is reasonable to assume that the noise $\mathbf{n}_i$ at different sensors are mutually uncorrelated. Here we consider the coherent multiple access channels (MAC), which means the data from different sensors are superimposed at the fusion center. Here we assume that the transmissions is the network are time-synchronous, i.e. the FC receives data from different sensors in the same time slot, which can be realized via GPS system. The collected data at the fusion center is corrupted by additive Gaussian noise $\mathbf{n}_0$. Without loss of generality, $\mathbf{n}_0$ is assumed to be white and zero-mean, i.e. $\mathbf{n}_0 \in \mathbb{C}^{M\times 1} \sim \mathcal{CN}(\mathbf{0}, \sigma_0^2\mathbf{I}_M)$.

Based on the system model above, the transmitted signal at the i-th sensor is $\mathbf{F}_i(\mathbf{s} +$

$\mathbf{n}_i$), and the received signal at the fusion center is presented as:

$$\mathbf{r} = \sum_{i=1}^{L} \mathbf{H}_i \mathbf{F}_i \big(\mathbf{s} + \mathbf{n}_i\big) + \mathbf{n}_0 \tag{6.1}$$

$$= \Big(\sum_{i=1}^{L} \mathbf{H}_i \mathbf{F}_i\Big)\mathbf{s} + \Big(\underbrace{\sum_{i=1}^{L} \mathbf{H}_i \mathbf{F}_i \mathbf{n}_i + \mathbf{n}_0}_{\mathbf{n}}\Big), \tag{6.2}$$

where the compound noise vector $\mathbf{n}$ is still Gaussian, i.e. $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_n})$ with its covariance matrix $\mathbf{\Sigma_n}$ as

$$\mathbf{\Sigma_n} = \sigma_0^2 \mathbf{I}_M + \sum_{i=1}^{L} \mathbf{H}_i \mathbf{F}_i \mathbf{\Sigma}_i \mathbf{F}_i^H \mathbf{H}_i^H. \tag{6.3}$$

It should be pointed out that the whiteness assumption of the Gaussian noise $\mathbf{n}_0$ at the receiver does not undermine generality of the model. Indeed if $\mathbf{n}_0 \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma}_0)$ has coloured covariance $\mathbf{\Sigma}_0$, by redefining $\widetilde{\mathbf{r}} \triangleq \mathbf{\Sigma}_0^{-\frac{1}{2}} \mathbf{r}$, $\widetilde{\mathbf{H}}_i \triangleq \mathbf{\Sigma}_0^{-\frac{1}{2}} \mathbf{H}_i$ and $\widetilde{\mathbf{n}}_0 \triangleq \mathbf{\Sigma}_0^{-\frac{1}{2}} \mathbf{n}_0$, the received signal can be equivalently written as

$$\widetilde{\mathbf{r}} = \sum_{i=1}^{L} \widetilde{\mathbf{H}}_i \mathbf{F}_i \big(\mathbf{s} + \mathbf{n}_i\big) + \widetilde{\mathbf{n}}_0, \tag{6.4}$$

with $\widetilde{\mathbf{n}}_\mathbf{0} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$, which coincides with the model in (6.1).

The mutual information between the source signal and the received signal at FC can

be given as

$$\begin{aligned}&\mathsf{MI}\Big(\big\{\mathbf{F}_i\big\}_{i=1}^{L}\Big)\\&=\log\det\left\{\mathbf{I}_M+\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\Big)\boldsymbol{\Sigma}_{\mathbf{s}}\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\Big)^H\Big(\sigma_0^2\mathbf{I}+\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\Big)^{-1}\right\}\end{aligned}\tag{6.5}$$

In practice, each sensor has independent transmission power according to its own battery condition. The average transmitted power for the i-th sensor is $\mathsf{E}\big\{\big\|\mathbf{F}_i\big(\mathbf{s}+\mathbf{n}_i\big)\big\|_2^2\big\}=\mathsf{Tr}\big\{\mathbf{F}_i\big(\boldsymbol{\Sigma}_{\mathbf{s}}+\boldsymbol{\Sigma}_i\big)\mathbf{F}_i^H\big\}$, which must respect its power constraint P_i. Thus the beamforming problem of the multiple sensor system can be formulated as the following optimization problem:

$$(\mathsf{P}0):\max.\mathsf{MI}\big(\{\mathbf{F}_i\}_i^L\big),\tag{6.6a}$$

$$\text{s.t.}\,\mathsf{Tr}\big\{\mathbf{F}_i\big(\boldsymbol{\Sigma}_{\mathbf{s}}+\boldsymbol{\Sigma}_i\big)\mathbf{F}_i^H\big\}\leq P_i,\ i\in\{1,\cdots,L\}.\tag{6.6b}$$

The above optimization problem is nonconvex, which can be easily seen by examining the convexity of the special case where $\{\mathbf{F}_i\}_{i=1}^L$ are all scalars. Efficient solutions to $(\mathsf{P}0)$ are desirable.

Since the above problem can hardly be solved in one shot, we propose iterative algorithms which fall in the framework of block coordinate descent/ascent (BCD/A) algorithms [63], also known as alternative minimization/maximization algorithm(AMA) [62] or Gauss-Seidel (GS) algorithms somewhere else [63] [64].

6.3 Algorithm Design

In this section, we focus on solutions to the problem (P0). Note that directly utilizing BCA method to partition the beamformers into groups does not help to simplify our problem. Even if only one separate beaformer is considered, the objective is still hard. Inspired by the weighted mean square error(WMMSE) method proposed by the seminal papers [73] and [60], we introduce auxiliary variables to convert the objective into a BCA-friendly form and then decompose the problem into solvable subproblems. Interestingly, although mutual information is independent of processing techniques at the receiver, our solution actually introduces a *virtual* linear filter at the fusion center to achieve our goal.

Firstly we introduce two useful lemmas which pave the way for transforming the original hard problem (P0).

Lemma 6.3.1 ([60, 73])**.** For any positive definite matrix $\mathbf{E} \in \mathcal{H}_{++}^{n}$, the following fact holds true

$$-\log\det(\mathbf{E}) = \max_{\mathbf{W}\in\mathcal{H}_{++}^{n}} \Big\{ \log\det\big(\mathbf{W}\big) - \mathsf{Tr}\{\mathbf{W}\mathbf{E}\} + n \Big\} \tag{6.7}$$

with the optimal solution $\mathbf{W}^{\star}$ given as

$$\mathbf{W}^{\star} = \mathbf{E}^{-1}. \tag{6.8}$$

Lemma 6.3.2. Define a matrix function $\mathbf{E}\big(\mathbf{G}\big)$ of variable $\mathbf{G}$ as

$$\mathbf{E}\big(\mathbf{G}\big) \triangleq \big(\mathbf{I} - \mathbf{G}^{H}\mathbf{H}\big)\boldsymbol{\Sigma}_{\mathbf{s}}\big(\mathbf{I} - \mathbf{G}^{H}\mathbf{H}\big)^{H} + \mathbf{G}^{H}\boldsymbol{\Sigma}_{\mathbf{n}}\mathbf{G}, \tag{6.9}$$

with $\mathbf{\Sigma_s}$ and $\mathbf{\Sigma_n}$ being positive definite matrices. Then for any positive definite matrix $\mathbf{W}$, the following optimization problem

$$\min_{\mathbf{G}} . \mathsf{Tr}\big\{\mathbf{W}\mathbf{E}(\mathbf{G})\big\} \tag{6.10}$$

can be solved by the optimal solution

$$\mathbf{G}^{\star} = \big(\mathbf{H}\mathbf{\Sigma_s}\mathbf{H}^H + \mathbf{\Sigma_n}\big)^{-1}\mathbf{H}\mathbf{\Sigma_s}. \tag{6.11}$$

At the same time, $\mathbf{E}(\mathbf{G}^{\star})$ is given as

$$\mathbf{E}(\mathbf{G}^{\star}) = \big(\mathbf{H}^H\mathbf{\Sigma_n}^{-1}\mathbf{H} + \mathbf{\Sigma_s}^{-1}\big)^{-1}. \tag{6.12}$$

Proof. The problem in (6.10) is a convex problem. To see this, notice that the objective function in (6.10) is a quadratic function of $\mathbf{G}$ with its quadratic terms being given as

$$\mathsf{Tr}\big\{\mathbf{W}\mathbf{G}^H\mathbf{H}\mathbf{\Sigma_s}\mathbf{H}^H\mathbf{G}\big\} + \mathsf{Tr}\big\{\mathbf{W}\mathbf{G}^H\mathbf{\Sigma_n}\mathbf{G}\big\}. \tag{6.13}$$

By the identities $\mathsf{Tr}\{\mathbf{AB}\} = \mathsf{Tr}\{\mathbf{BA}\}$ and $\mathsf{Tr}\{\mathbf{ABCD}\} = vec^T(\mathbf{D}^T)\big[\mathbf{C}^T \otimes \mathbf{A}\big]vec(\mathbf{B})$, the first term of the above quadratic terms can be rewritten as

$$\mathsf{Tr}\big\{\mathbf{W}\mathbf{G}^H\mathbf{H}\mathbf{\Sigma_s}\mathbf{H}^H\mathbf{G}\big\} = vec^H(\mathbf{G})\Big[\mathbf{W}^* \otimes \big(\mathbf{H}\mathbf{\Sigma_s}\mathbf{H}^H\big)\Big]vec(\mathbf{G}). \tag{6.14}$$

Notice that $\mathbf{W}$ and $\mathbf{H}\mathbf{\Sigma_s}\mathbf{H}^H$ are both positive semi-definite, so $\Big[\mathbf{W}^* \otimes \big(\mathbf{H}\mathbf{\Sigma_s}\mathbf{H}^H\big)\Big]$ is positive semi-definite [67] and thus the first quadratic term is a convex function of $\mathbf{G}$. Similarly the second quadratic term in (6.13) can also be proved to be convex function of $\mathbf{G}$. Thus (6.10) is non-constrained convex problem of $\mathbf{G}$. By setting the derivative

with respective to $\mathbf{G}$ to zero [78], we obtain

$$\frac{\partial \mathsf{Tr}\{\mathbf{WE}(\mathbf{G})\}}{\partial \mathbf{G}^*} = \Big[\Big(\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{H}^H + \boldsymbol{\Sigma}_{\mathbf{n}}\Big)\mathbf{G} - \mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\Big]\mathbf{W} = \mathbf{O}. \tag{6.15}$$

Notice that $\mathbf{W}$ is positive definite, it can be cancelled and thus the equation (6.11) has been obtained. By substituting (6.11) into (6.9), (6.12) can be proved. □

Comment 6.3.1. For the special case $\mathbf{W} = \mathbf{I}$, the result in lemma 6.3.2 is the well known Wiener filter. Here lemma 6.3.2 actually slightly generalizes this well known result. As we have shown above, when the mean square error is weighted by a matrix $\mathbf{W}$, the Wiener filter maintains its optimality as long as the weighted parameter $\mathbf{W}$ is positive definite.

Now by introducing the notation

$$\widetilde{\mathbf{H}} \triangleq \sum_{i=1}^{L} \mathbf{H}_i \mathbf{F}_i, \tag{6.16}$$

and the notations in equation (6.3), we can transform our objective function $\mathsf{MI}\big(\{\mathbf{F}_i\}_{i=1}^{L}\big)$

as the following:

$$\mathsf{MI}\Big(\{\mathbf{F}_i\}_{i=1}^{L}\Big)=\log\det\Big(\mathbf{I}_M+\widetilde{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\widetilde{\mathbf{H}}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\Big) \tag{6.17}$$

$$=\log\det\Big(\big(\widetilde{\mathbf{H}}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\widetilde{\mathbf{H}}+\boldsymbol{\Sigma}_{\mathbf{s}}^{-1}\big)\boldsymbol{\Sigma}_{\mathbf{s}}\Big) \tag{6.18}$$

$$=-\log\det\Big(\widetilde{\mathbf{H}}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\widetilde{\mathbf{H}}+\boldsymbol{\Sigma}_{\mathbf{s}}^{-1}\Big)^{-1}+\log\det\big(\boldsymbol{\Sigma}_{\mathbf{s}}\big) \tag{6.19}$$

$$=\max_{\mathbf{W}\in\mathcal{H}_{++}^{K}}\Big\{\log\det\big(\mathbf{W}\big)-\mathsf{Tr}\big\{\mathbf{W}\big(\widetilde{\mathbf{H}}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\widetilde{\mathbf{H}}+\boldsymbol{\Sigma}_{\mathbf{s}}^{-1}\big)^{-1}\big\}+K\Big\}+\log\det\big(\boldsymbol{\Sigma}_{\mathbf{s}}\big) \tag{6.20}$$

$$=\max_{\substack{\mathbf{W}\in\mathcal{H}_{++}^{K},\\ \mathbf{G}}}\Big\{\log\det\big(\mathbf{W}\big)-\mathsf{Tr}\Big\{\mathbf{W}\big[\big(\mathbf{I}-\mathbf{G}^H\widetilde{\mathbf{H}}\big)\boldsymbol{\Sigma}_{\mathbf{s}}\big(\mathbf{I}-\mathbf{G}^H\widetilde{\mathbf{H}}\big)^H+\mathbf{G}^H\boldsymbol{\Sigma}_{\mathbf{n}}\mathbf{G}\big]\Big\}\Big\}$$

$$+K+\log\det\big(\boldsymbol{\Sigma}_{\mathbf{s}}\big), \tag{6.21}$$

where the last two steps follow lemma 6.3.1 and 6.3.2 respectively.

Thus the optimization problem (P0) maximizing MI has been transformed into an equivalent problem (P1) in (6.22) as follows

$$(\mathsf{P1})\ \max_{\substack{\mathbf{W}\in\mathcal{H}_{++}^{K},\\ \{\mathbf{F}_i\}_{i=1}^{L},\mathbf{G}}}\ \mathsf{MI}\Big(\{\mathbf{F}_i\}_{i=1}^{L},\mathbf{W},\mathbf{G}\Big) \tag{6.22a}$$

$$=\Big\{\log\det\big(\mathbf{W}\big)-\mathsf{Tr}\Big\{\mathbf{W}\Big[\Big(\mathbf{I}-\mathbf{G}^H\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\Big)\boldsymbol{\Sigma}_{\mathbf{s}}\Big(\mathbf{I}-\mathbf{G}^H\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\Big)^H+\mathbf{G}^H\boldsymbol{\Sigma}_{\mathbf{n}}\mathbf{G}\Big]\Big\}\Big\}$$

$$+\log\det(\boldsymbol{\Sigma}_{\mathbf{s}})+K,$$

$$\text{s.t. }\mathsf{Tr}\big\{\mathbf{F}_i\big(\boldsymbol{\Sigma}_{\mathbf{s}}+\boldsymbol{\Sigma}_{\mathbf{n}}\big)\mathbf{F}_i^H\big\}\le P_i,\qquad i\in\{1,\cdots,L\}. \tag{6.22b}$$

As a straightforward consequence of the above two lemmas, we have obtained the optimal solutions to the following two subproblems of (P1).

When $\{\mathbf{F}_i\}_{i=1}^L$ and $\mathbf{G}$ are given, the optimal $\mathbf{W}^\star$ is given as

$$\begin{aligned}\mathbf{W}^\star &= \arg\max_{\mathbf{W}\in\mathcal{H}_{++}^K} \mathsf{MI}\Big(\mathbf{W}\Big|\{\mathbf{F}_i\}_{i=1}^L, \mathbf{G}\Big) \\ &= \Big[\Big(\mathbf{I}–\mathbf{G}^H\big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\big)\Big)\mathbf{\Sigma_s}\Big(\mathbf{I}–\mathbf{G}^H\big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\big)\Big)^H + \mathbf{G}^H\mathbf{\Sigma_n}\mathbf{G}\Big]^{-1}.\end{aligned} \tag{6.23}$$

When $\{\mathbf{F}_i\}_{i=1}^L$ and $\mathbf{W}$ are given, the optimal $\mathbf{G}^\star$ is given as

$$\begin{aligned}\mathbf{G}^\star &= \arg\max_{\mathbf{G}} .\mathsf{MI}\Big(\mathbf{G}\Big|\{\mathbf{F}_i\}_{i=1}^L, \mathbf{W}\Big) \\ &= \Big[\big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\big)\mathbf{\Sigma_s}\big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\big)^H + \mathbf{\Sigma_n}\Big]^{-1}\Big(\sum_{i=1}^L \mathbf{H}_i\mathbf{F}_i\Big)\mathbf{\Sigma_s},\end{aligned} \tag{6.24}$$

with $\mathbf{\Sigma_n}$ being given in equation (6.3).

Now we focus on the subproblem of optimizing $\{\mathbf{F}_i\}_{i=1}^L$ with $\mathbf{W}$ and $\mathbf{G}$ given. Towards this end, we have two options—we can either jointly optimize $\{\mathbf{F}_i\}_{i=1}^L$ in one shot, or we can further consult to BCA methodology again to partition the entire variables $\{\mathbf{F}_i\}_{i=1}^L$ into L blocks, $\{\mathbf{F}_1\}$, $\cdots$, $\{\mathbf{F}_L\}$ and attack L smaller problems one by one in a cyclic manner. For both of these two options, solutions, hopefully in a closed form, are desirable and complexity are concerned. In the following, we discuss these two alternatives in details.

6.3.1 Jointly Optimizing $\{\mathbf{F}_i\}_{i=1}^{L}$

The subproblem of (P1) maximizing $\mathsf{MI}\big(\{\mathbf{F}_i\}_{i=1}^{L}\big|\mathbf{W},\mathbf{G}\big)$ with $\mathbf{W}$ and $\mathbf{G}$ given is rewritten as follows

$$(\mathsf{P2})\ \min_{\{\mathbf{F}_i\}_{i=1}^{L}}.\mathsf{Tr}\Big\{\mathbf{W}\Big[\Big(\mathbf{I}\text{–}\mathbf{G}^H\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\Big)\mathbf{\Sigma_s}\Big(\mathbf{I}\text{–}\mathbf{G}^H\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\Big)^H+\mathbf{G}^H\mathbf{\Sigma_n}\mathbf{G}\Big]\Big\}, \quad (6.25a)$$

$$\text{s.t. } \mathsf{Tr}\big\{\mathbf{F}_i\big(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\big)\mathbf{F}_i^H\big\}\leq P_i, \qquad i\in\{1,\cdots,L\}. \quad (6.25b)$$

The following theorem identifies the convexity of (P2).

Theorem 6.3.1. The problem $(P2)$ is convex.

Proof. To begin with, we first look at the function $f\big(\mathbf{X}\big):\mathbb{C}^{m\times n}\mapsto\mathbb{R}$ given as follows:

$$f\big(\mathbf{X}\big)\triangleq\mathsf{Tr}\big\{\mathbf{\Sigma}_1\mathbf{X}\mathbf{\Sigma}_2\mathbf{X}^H\big\} \quad (6.26)$$

with constant matrices $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ being positive semi-definite and having appropriate dimensions. By the identity $\mathsf{Tr}\{\mathbf{ABCD}\}=vec^T(\mathbf{D}^T)\big[\mathbf{C}^T\otimes\mathbf{A}\big]vec(\mathbf{B})$, $f\big(\mathbf{X}\big)$ can be equivalently written as

$$f\big(\mathbf{X}\big)=vec^H\big(\mathbf{X}\big)\big[\mathbf{\Sigma}_2^*\otimes\mathbf{\Sigma}_1\big]vec\big(\mathbf{X}\big). \quad (6.27)$$

Since $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are positive semi-definite, $[\mathbf{\Sigma}_1^*\otimes\mathbf{\Sigma}_2]$ is positive semi-definite [67]. Thus $f\big(\mathbf{X}\big)$ is actually a convex function with respect to $\mathbf{X}$.

For a further step, we replace $\mathbf{X}=\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i$. Since $\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i$ is an affine (linear actually) transformation of variables $\{\mathbf{F}_i\}_{i=1}^{L}$, and affine operations preserve convexity

by [68], the following function

$$f\big(\{\mathbf{F}_i\}_{i=1}^{L}\big) = \mathsf{Tr}\Big\{\boldsymbol{\Sigma}_1\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\boldsymbol{\Sigma}_2\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)^H\Big\} \tag{6.28}$$

is a convex function with respect to variables $\{\mathbf{F}_i\}_{i=1}^{L}$ jointly.

To identify the convexity of the objective in (6.25a), it suffices to prove the nonlinear terms of $\{\mathbf{F}_i\}_{i=1}^{L}$ are convex, which are given as

$$\mathsf{Tr}\Big\{\big(\mathbf{G}\mathbf{W}\mathbf{G}^H\big)\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\boldsymbol{\Sigma}_{\mathbf{s}}\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)^H\Big\} + \sum_{i=1}^{L}\mathsf{Tr}\Big\{\big(\mathbf{H}_i^H\mathbf{G}\mathbf{W}\mathbf{G}^H\mathbf{H}_i\big)\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\Big\}. \tag{6.29}$$

Based on the discussion at the beginning of this proof, each of above terms is convex and thus the objective is convex. Similarly the convexity of each power constraint function can also be recognized. Thus the problem (P2) is convex. □

After identifying the convexity of problem (P2), we reformulate it into a standard quadratic constrained quadratic problem(QCQP) problem. To this end, we introduce the

following notations

$$\mathbf{f}_i \triangleq vec\big(\mathbf{F}_i\big); \tag{6.30a}$$

$$\mathbf{g} \triangleq vec\big(\mathbf{G}\big); \tag{6.30b}$$

$$\mathbf{A}_{ij} \triangleq \mathbf{\Sigma}_{\mathbf{s}}^{*} \otimes \Big(\mathbf{H}_i^H \mathbf{G}\mathbf{W}\mathbf{G}^H \mathbf{H}_j\Big); \tag{6.30c}$$

$$\mathbf{B}_i \triangleq \big(\mathbf{W}\mathbf{\Sigma}_{\mathbf{s}}\big)^{*} \otimes \mathbf{H}_i; \tag{6.30d}$$

$$\mathbf{C}_i \triangleq \mathbf{\Sigma}_i^{*} \otimes \Big(\mathbf{H}_i^H \mathbf{G}\mathbf{W}\mathbf{G}^H \mathbf{H}_i\Big); \tag{6.30e}$$

$$\mathbf{f} \triangleq \big[\mathbf{f}_1^T, \cdots, \mathbf{f}_L^T\big]^T; \tag{6.30f}$$

$$\mathbf{A} \triangleq \big[\mathbf{A}_{ij}\big]_{i,j=1}^{L}; \tag{6.30g}$$

$$\mathbf{B} \triangleq \big[\mathbf{B}_1, \cdots, \mathbf{B}_L\big]; \tag{6.30h}$$

$$\mathbf{C} \triangleq \mathsf{Diag}\big\{\mathbf{C}_1, \cdots, \mathbf{C}_L\big\}; \tag{6.30i}$$

$$\mathbf{D}_i \triangleq \mathsf{Diag}\Big\{\mathbf{O}_{K(\sum_{j=1}^{i-1} N_j)}, \big(\mathbf{\Sigma}_{\mathbf{s}} + \mathbf{\Sigma}_i\big)^{*} \otimes \mathbf{I}_{N_i}, \mathbf{O}_{K(\sum_{j=i+1}^{L} N_j)}\Big\}; \tag{6.30j}$$

$$c \triangleq \mathsf{Tr}\big\{\mathbf{W}\mathbf{\Sigma}_{\mathbf{s}}\big\} + \sigma_0^2 \mathsf{Tr}\big\{\mathbf{G}\mathbf{W}\mathbf{G}^H\big\}. \tag{6.30k}$$

Based on the above notations, problem (P2) can be equivalently written as the following QCQP problem,

$$(\mathsf{P3}): \min_{\mathbf{f}} \ \mathbf{f}^H\big(\mathbf{A}+\mathbf{C}\big)\mathbf{f} - 2\mathsf{Re}\big\{\mathbf{g}^H\mathbf{B}\mathbf{f}\big\} + c, \tag{6.31a}$$

$$s.t. \ \ \mathbf{f}^H\mathbf{D}_i\mathbf{f} \le P_i, \qquad i \in \{1, \cdots, L\}. \tag{6.31b}$$

By theorem 1, (P3) is convex, thus $(\mathbf{A}+\mathbf{C})$ is positive semidefinite, which implies that its square root $(\mathbf{A}+\mathbf{C})^{\frac{1}{2}}$ exists. Therefore the above problem can be further rewritten in

a standard SOCP form as follows:

$$(\mathsf{P3}_{SOCP}) : \min_{\mathbf{f},t,s} . \ t, \tag{6.32a}$$

$$\text{s.t. } s - 2\mathsf{Re}\{\mathbf{g}^H\mathbf{B}\mathbf{f}\} + c \leq t; \tag{6.32b}$$

$$\left\| \begin{array}{c} (\mathbf{A}+\mathbf{C})^{\frac{1}{2}}\mathbf{f} \\ \frac{s-1}{2} \end{array} \right\|_2 \leq \frac{s+1}{2}; \tag{6.32c}$$

$$\left\| \begin{array}{c} \mathbf{D}_i^{\frac{1}{2}}\mathbf{f} \\ \frac{P_i-1}{2} \end{array} \right\|_2 \leq \frac{P_i+1}{2}, \ \ i \in \{1, \cdots, L\}. \tag{6.32d}$$

The above problem can be solved by standard numerical tools like CVX [69].

The method discussed above is summarized in algorithm 4.

Algorithm 4: 3-Block BCA Algorithm to solve (P0)

Initialization: randomly generate feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^L$; obtain $\mathbf{G}^{(0)}$ by (6.24); obtain $\mathbf{W}^{(0)}$ by (6.23);
repeat
  with $\mathbf{G}^{(j-1)}$ and $\mathbf{W}^{(j-1)}$ being fixed, solve (P3) in (6.32), obtain $\{\mathbf{F}_i^{(j)}\}_{i=1}^L$;
  with $\{\mathbf{F}_i^{(j)}\}_{i=1}^L$ and $\mathbf{W}^{(j-1)}$ being fixed, obtain $\mathbf{G}^{(j)}$ by (6.24);
  with $\{\mathbf{F}_i^{(j)}\}_{i=1}^L$ and $\mathbf{G}^{(j)}$ being fixed, obtain $\mathbf{W}^{(j)}$ by (6.23);
until *increase of* MI *is sufficiently small or predefined number of iterations is reached*;

For the proposed 3-BCA algorithm, we have the following conclusion on its convergence

Theorem 6.3.2. Assume that the covariance matrix $\mathbf{\Sigma_s} \succ 0$. Algorithm 4 generates increasing MI sequence. Its solution sequence has limit points, and each limit point of the solution sequence is a KKT point of the original problem $(P0)$.

Proof. Refer to appendix 6.6.1. □

6.3.2 Cyclic $(L+1)$-BCA Algorithm

Although the above proposed 3-block BCA algorithm guarantees a satisfactory convergence, the subproblem (P3) relies on standard numerical solvers, e.g. interior point method [69], to obtain solutions. Closed form solutions to (P3) is unknown. According to the complexity analysis performed in next subsection 6.3.3, when the number of sensors and/or antenna number of each sensor grows, the problem (P3) can be very large size and consequently highly computation demanding. So effective algorithms with lower complexity are desirable. In this subsection, we consult to BCA methodology again to further partition the variables $\{\mathbf{F}_i\}_{i=1}^{L}$ into L singleton sets: $\{\mathbf{F}_1\},\cdots,\{\mathbf{F}_L\}$. This results in a cyclic $(L+1)$-BCA algorithm, where only one separate beamformer $\mathbf{F}_i$ is optimized at each time and different beamformers are updated in an round robin manner.

Now the problem of updating one separate beamformer becomes

$$(\mathsf{P}3_i)\ \min_{\mathbf{f}_i}.\ \mathbf{f}_i^H\big(\mathbf{A}_{ii}+\mathbf{C}_i\big)\mathbf{f}_i - 2\mathsf{Re}\big\{\big(\mathbf{g}^H\mathbf{B}_i-\mathbf{q}_i^H\big)\mathbf{f}_i\big\}, \tag{6.33a}$$

$$\text{s.t. } \mathbf{f}_i^H\mathbf{E}_i\mathbf{f}_i \le P_i \tag{6.33b}$$

with the definitions of $\mathbf{q}_i$ and $\mathbf{E}_i$ as follows

$$\mathbf{q}_i \triangleq \sum_{j\neq i}\mathbf{A}_{ij}\mathbf{f}_j;\ \ \mathbf{E}_i \triangleq \big(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\big)^*\otimes\mathbf{I}_{N_i}. \tag{6.34}$$

We introduce the following notations

$$\mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{A}_{ii}+\mathbf{C}_i\big)\mathbf{E}_i^{-\frac{1}{2}} = \mathbf{U}_i \begin{bmatrix} \lambda_{i,1} & & \\ & \ddots & \\ & & \lambda_{i,KN_i} \end{bmatrix} \mathbf{U}_i^H; \tag{6.35a}$$

$$\mathbf{p}_i = \mathbf{U}_i^H \mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{B}_i^H\mathbf{g} - \mathbf{q}_i\big), \tag{6.35b}$$

with the eigenvalues $\{\lambda_{i,j}\}_{j=1}^{KN_i}$ arranged in an decreasing order, i.e. $\lambda_{i,1} \geq \cdots \geq \lambda_{i,KN_i}$. We denote the k-th element of $\mathbf{p}_i$ as $p_{i,k}$ and assume that $r_i = \mathsf{rank}\big(\mathbf{A}_{ii}+\mathbf{C}_i\big)$.

Then the solution to problem ($\mathsf{P}3_i$) is given by the following theorem.

Theorem 6.3.3. Under the assumption that $\mathbf{\Sigma_s} \succ 0$ or $\mathbf{\Sigma}_i \succ 0$, $i \in \{1, \cdots, L\}$, the optimal solution of problem ($P3_i$) is given as follows:

CASE(I)—if either of the following two conditions holds:

i) $\exists k \in \{r_i+1, \cdots, KN_i\}$ such that $|p_{i,k}| \neq 0$;

or ii) $\sum_{k=r_i+1}^{KN_i} |p_{i,k}| = 0$ and $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} > P_i$.

The optimal solution to ($P3_i$) is given by

$$\mathbf{f}_i^\star = \big(\mathbf{A}_{ii}+\mathbf{C}_i+\mu_i^\star\mathbf{E}_i\big)^{-1}\big(\mathbf{B}_i^H\mathbf{g} - \mathbf{q}_i\big), \tag{6.36}$$

with the positive value $\mu_i^\star$ being the unique solution to the following equation:

$$f(\mu_i) = \sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{(\lambda_{i,k}+\mu_i)^2} = P_i. \tag{6.37}$$

CASE(II)—otherwise,i.e.$\sum_{k=r_i+1}^{KN_i} |p_{i,k}| = 0$,$\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} \leq P_i$,

The optimal solution to $(P3_i)$ is given by

$$\mathbf{f}_i^\star = \mathbf{E}_i^{-\frac{1}{2}}\Big(\mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{A}_{ii}+\mathbf{C}_i\big)\mathbf{E}_i^{-\frac{1}{2}}\Big)^\dagger \mathbf{E}_i^{-\frac{1}{2}}\big(\mathbf{B}_i^H\mathbf{g}-\mathbf{q}_i\big). \tag{6.38}$$

Proof. For limit of space, please refer to Theorem 3 in [71] for detailed proof. □

In CASE(I) of theorem 6.3.3, equation (6.37) generally has no closed form solution. Notice that $f(\mu_i)$ is a one-dimension strictly decreasing function of μ_i. So the determination of $\mu_i^\star$ can be efficiently performed by a bisection search. Thus a finite interval containing $\mu_i^\star$ is necessary from which the bisection search can start. The following lemma provides us bounds for $\mu_i^\star$.

Lemma 6.3.3. A lower bound lbd_i and upper bound ubd_i for positive $\mu_i^\star$ in (6.36) in theorem 6.3.3 can be given as follows:

i) For subcase i) of CASE(I)

$$lbd_i = \left[\frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,1}\right]^+, \quad ubd_i = \frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}}; \tag{6.39}$$

ii) For subcase ii) of CASE(I)

$$lbd_i = \left[\frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,1}\right]^+, \quad ubd_i = \frac{\|\mathbf{p}_i\|_2}{\sqrt{P_i}} - \lambda_{i,r_i}, \tag{6.40}$$

where $[x]^+ \triangleq \max(x, 0)$.

Proof. For limit of space, please refer to Lemma 1 in [71] for detailed proof. □

By theorem 6.3.3 and lemma 6.3.3, we have obtained a nearly closed form solution

to the problem ($\mathsf{P}3_i$). Here we claim the above solution *nearly* closed form since it involves a bisection search.

It is worth noting that the *fully* closed form solution to problem ($\mathsf{P}3_i$) does exist in the special but important case of scalar source signal, i.e. $K = 1$. As will be seen in the complexity analysis in subsection 6.3.3, the increase of source signal dimension can extensively enlarge the beamforming problem size and therefore its complexity. So in practice, when the wireless sensor network has adequate bandwidth, it is preferred to transmit the sensed data component by component to decrease the processing and beamforming design complexity at the fusion center. Another tempting reason to do so is that when the source signal is scalar, $\mu_i^\star$ can be obtained in an explicit way and therefore the bisection search is not needed. Moreover, at this time, the eigenvalue decomposition (recall (6.35a)) will not be involved in solving ($\mathsf{P}3_i$). This conclusion reads as the following corollary.

Corollary 6.3.1. For scalar transmission ($K = 1$), fully analytic solution to subproblem ($P3_i$) can be obtained without evoking bisection search or eigenvalue decomposition.

Proof. In the special case where the signal source is scalar, the variables and parameters in the subproblem optimizing one separate beamformer are specialized as follows

$$\mathbf{W} \to w;\ \mathbf{F}_i \to \mathbf{f}_i;\ \mathbf{G} \to \mathbf{g};\ \mathbf{\Sigma_s} \to \sigma_s^2;\ \mathbf{\Sigma}_i \to \sigma_i^2. \tag{6.41}$$

By defining $\widetilde{\mathbf{q}}_i \triangleq \sum_{j\neq i} \mathbf{H}_j\mathbf{f}_j$, ignoring the terms independent of $\mathbf{f}_i$ and omitting the

constant positive factor w in the objective, the problem $(\mathsf{P3}_i)$ is rewritten as follows

$$(\mathsf{P3}_i): \min_{\mathbf{f}_i} . \ (\sigma_s^2+\sigma_i^2)\mathbf{f}_i^H\big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big)\mathbf{f}_i$$
$$-2\sigma_s^2\mathsf{Re}\big\{(1-\widetilde{\mathbf{q}}_i^H\mathbf{g})\mathbf{g}^H\mathbf{H}_i\mathbf{f}_i\big\} \tag{6.42a}$$
$$\text{s.t. } \|\mathbf{f}_i\|^2 \le \frac{P_i}{\sigma_s^2+\sigma_i^2} \triangleq \bar{P}_i. \tag{6.42b}$$

Solving the problem $(\mathsf{P3}_i)$ just follows the outline of theorem 6.3.3. Here, the key point leading to a closed form solution is the fact that the quadratic matrix $\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i$ has rank-1, i.e. $r_i = 1$ in theorem 6.3.3. Thus we obtain

$$(\sigma_s^2+\sigma_i^2)\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i \tag{6.43}$$
$$= \mathbf{U}_i \begin{bmatrix} (\sigma_s^2+\sigma_i^2)\mathbf{g}^H\mathbf{H}_i\mathbf{H}_i^H\mathbf{g} & \mathbf{0}^H \\ \mathbf{0} & \mathbf{O}_{(KN_i-1)\times(KN_i-1)} \end{bmatrix} \mathbf{U}_i^H,$$

with unitary matrix $\mathbf{U}_i \triangleq \big[\mathbf{u}_{i,1}, \mathbf{u}_{i,2}, \cdots, \mathbf{u}_{i,KN_i}\big]$ having its columns $\{\mathbf{u}_{i,j}\}_{j=1}^{KN_i}$ satisfying the following properties

$$\mathbf{u}_{i,1} = \frac{\mathbf{H}_i^H\mathbf{g}}{\|\mathbf{H}_i^H\mathbf{g}\|_2}, \text{ and } \mathbf{u}_{i,j}^H\mathbf{H}_i^H\mathbf{g} = 0, \text{ for } j = 2, \cdots, KN_i. \tag{6.44}$$

It can be readily checked that the parameter $\mathbf{p}_i$ in theorem 6.3.3 is given as:

$$p_{i,1} = \sigma_s^2(1-\mathbf{g}^H\widetilde{\mathbf{q}}_i)\|\mathbf{H}_i^H\mathbf{g}\|_2; \ p_{i,j} = 0, j = \{2, \cdots, KN_i\}. \tag{6.45}$$

At this time, the function $f(\mu_i)$ in (6.37) reduces to an elegant form

$$f(\mu_i) = \frac{\sigma_s^4 \left|1 - \mathbf{g}^H \widetilde{\mathbf{q}}_i\right|^2 \left\|\mathbf{H}_i^H \mathbf{g}\right\|_2^2}{\left(\mu_i + (\sigma_s^2 + \sigma_i^2)\mathbf{g}^H \mathbf{H}_i \mathbf{H}_i^H \mathbf{g}\right)^2}. \tag{6.46}$$

Based on the above observations, it can be concluded that the subcase $i)$ of CASE(I) in theorem 6.3.3 will never occur. The two cases for positive and zero $\mu_i^\star$ can be specified as follows:

CASE(I)— $\mu_i^\star > 0$

This is equivalent to $\sigma_s^4 \left|1 - \mathbf{g}^H \widetilde{\mathbf{q}}_i\right|^2 > (\sigma_s^2 + \sigma_i^2)^2 \bar{P}_i \|\mathbf{H}_i^H \mathbf{g}\|_2^2$ and optimal solutions are determined by

$$\mu_i^\star = \sigma_s^2 \bar{P}_i^{-\frac{1}{2}} \left|1 - \mathbf{g}^H \widetilde{\mathbf{q}}_i\right| \left\|\mathbf{H}_i^H \mathbf{g}\right\|_2 - (\sigma_s^2 + \sigma_i^2) \left\|\mathbf{H}_i^H \mathbf{g}\right\|_2, \tag{6.47a}$$

$$\mathbf{f}_i^\star = \sigma_s^2 (1 - \mathbf{g}^H \widetilde{\mathbf{q}}_i) \left(\mu_i^\star \mathbf{I} + (\sigma_s^2 + \sigma_i^2) \mathbf{H}_i^H \mathbf{g} \mathbf{g}^H \mathbf{H}_i\right)^{-1} \mathbf{H}_i^H \mathbf{g}. \tag{6.47b}$$

CASE(II)— $\mu_i^\star = 0$

This holds if and only if $\sigma_s^4 \left|1 - \mathbf{g}^H \widetilde{\mathbf{q}}_i\right|^2 \leq (\sigma_s^2 + \sigma_i^2)^2 \bar{P}_i \|\mathbf{H}_i^H \mathbf{g}\|_2^2$ and the optimal $\mathbf{f}_i^\star$ is given by

$$\mathbf{f}_i^\star = \frac{\sigma_s^2 (1 - \mathbf{g}^H \widetilde{\mathbf{q}}_i) \mathbf{H}_i^H \mathbf{g}}{(\sigma_s^2 + \sigma_i^2) \mathbf{g}^H \mathbf{H}_i \mathbf{H}_i^H \mathbf{g}}. \tag{6.48}$$

Thus we have seen that for scalar transmission case, fully closed form solution to (P3$_i$) can be obtained without bisection search or eigenvalue decomposition. □

The cyclic $(L+1)$-BCA algorithm is summarized in algorithm 5.

Algorithm 5: Cyclic $(L+1)$-BCA Algorithm to Solve (P0) Optimizing MI

Initialization: randomly generate feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^{L}$; obtain $\mathbf{G}^{(0)}$ by (6.24); obtain $\mathbf{W}^{(0)}$ by (6.23);
repeat
  for $i = 1;\ i <= L;\ i++$ **do**
    with $\mathbf{G}$, $\mathbf{W}$, $\{\mathbf{F}_k\}_{k\neq i}$ being fixed, perform eigenvalue decomposition and obtain $\mathbf{p}_i$ (6.35);
    if $\exists k \in \{r_i+1, \cdots, KN_i\}$ *s.t.* $|p_{i,k}| \neq 0$ **then**
      determine bounds lbd_i and ubd_i by (6.39);
      bisection search on $\left[lbd_i, ubd_i\right]$ to determine $\mu_i^\star$ satisfying (6.37);
      obtain $\mathbf{F}_i$ by (6.36);
    else if $\sum_{k=r_i+1}^{KN_i} |p_{i,k}|^2 = 0$, $\sum_{k=1}^{r_i} \frac{|p_{i,k}|^2}{\lambda_{i,k}^2} > P_i$ **then**
      determine bounds lbd_i and ubd_i by (6.40);
      bisection search on $\left[lbd_i, ubd_i\right]$ to determine $\mu_i^\star$ satisfying (6.37);
      obtain $\mathbf{F}_i$ by (6.36);
    else
      obtain $\mathbf{F}_i$ by (6.38);
    end
    update $\mathbf{G}$ by (6.24) ;
    update $\mathbf{W}$ by (6.23) ;
  end
until *increase of* MI *is sufficiently small or predefined number of iterations is reached*;

6.3.3 Complexity

In this subsection, we discuss the complexity of the proposed algorithms.

The two subproblems optimizing $\mathbf{G}$ and $\mathbf{W}$ have closed form solutions in (6.24) and (6.23), their complexities come from matrix inversion and are given as $\mathcal{O}\big(K^3\big)$.

For the 3-block BCA algorithm, the SOCP problem ($\mathsf{P3}_{SOCP}$) in (6.32) is solved by jointly optimizing all beamformers. The complexity of solving an SOCP is [93]

$$\mathcal{O}\bigg(k_{SOC}^{\frac{1}{2}}\Big(m_{SOC}^3+m_{SOC}^2\sum_{i=1}^{k_{SOC}}n_{SOC,i}+\sum_{i=1}^{k_{SOC}}n_{SOC,i}^2\Big)\bigg), \tag{6.49}$$

where k_{SOC} is the number of second order cone constraints, m_{SOC} is the dimension of optimization problem and $n_{SOC,i}$ denotes the dimension of the i-th second order cone constraint. For the problem in (6.32), $k_{SOC}=L+1$, $m_{SOC}=K\big(\sum_{i=1}^{L}N_i\big)$, $n_{SOC,1}=K\big(\sum_{i=1}^{L}N_i\big)+1$ for the first second order cone constraint in (6.32c) and $n_{SOC,i+1}=KN_i+1$ for the i-th power constraint in (6.32d), $i\in\{1,\cdots,L\}$. Substituting these parameters into (6.49), the complexity of solving (P3) is $\mathcal{O}\Big(\sqrt{L}K^3(\sum_{i=1}^{L}N_i)^3\Big)$, this is also the complexity for each loop of 3-block BCA algorithm.

For the cyclic $(L+1)$-BCA algorithm, the problem ($\mathsf{P3}_i$) optimizing one separate sensor's beamformer has its major complexity coming from eigenvalue decomposition, which is $\mathcal{O}(K^3N_i^3)$. Thus the complexity for each loop is $\mathcal{O}\Big(\sum_{i=1}^{L}K^3N_i^3\Big)$. Clearly by fully decomposing the original problem and researching the solution structure of the subproblems, the $(L+1)$-block BCA algorithm effectively lowers the computation complexity.

6.4 Numerical Results

In this section, numerical results are presented to verify the algorithms proposed in the previous section.

In our following experiments, we test the case where the source signal and all observation noise are colored. Specifically, we set the covariance matrices of the source signal and observation noise as

$$\mathbf{\Sigma_s} = \sigma_s^2 \mathbf{\Sigma}_0, \quad \mathbf{\Sigma}_i = \sigma_i^2 \mathbf{\Sigma}_0, \quad i \in \{1, \cdots, L\}, \tag{6.50}$$

where the $K \times K$ Toeplitz matrix $\mathbf{\Sigma}_0$ is defined as

$$\mathbf{\Sigma}_0 = \begin{bmatrix} 1 & \rho & \rho^2 & \ddots & \rho^{K-1} \\ \rho & 1 & \rho & \ddots & \ddots \\ \rho^2 & \rho & 1 & \ddots & \rho^2 \\ \ddots & \ddots & \ddots & \ddots & \rho \\ \rho^{K-1} & \ddots & \rho^2 & \rho & 1 \end{bmatrix}. \tag{6.51}$$

The parameter ρ in the above equation is used to adjust the correlation level between different components of the signal or noise. In our test, ρ is set as $\rho = 0.5$. Here we define the observation signal to noise ratio at the i-th sensor as $\mathsf{SNR}_i \triangleq \frac{\sigma_s^2}{\sigma_i^2}$ and the channel signal to noise ratio as $\mathsf{SNR} \triangleq \frac{\sigma_s^2}{\sigma_0^2}$.

In figure 6.2 and 6.3 we test the performance of the 3-block BCA and cyclic $(L+1)$-BCA algorithms for multiple dimension source signal. Here two cases are tested—heterogeneous network and homogeneous network, in figures 6.2 and 6.3 respectively.

In the heterogeneous network, the transmission power, observation noise level and numbers of antennae of each sensor are different. We set up a wireless sensor network with three sensors, i.e. $L = 3$. The dimension of the source signal and the number of antenna of the FC are chosen as 3 and 4 respectively, i.e. $K = 3$ and $M = 4$. We randomly set the antenna number for each sensor as $N_1 = 3$, $N_2 = 4$, and $N_3 = 5$ respectively, the transmission power constraint for each sensor as $P_1 = 2$, $P_2 = 2$ and $P_3 = 3$ respectively and the observation signal to noise ratio for each sensor as $\mathsf{SNR}_1 = 8dB$, $\mathsf{SNR}_2 = 9dB$ and $\mathsf{SNR}_3 = 10dB$ respectively. Comparatively, in homogeneous sensor network each sensor has the same transmission power, observation noise level and number of antenna. In this test case we assume that $K = 4$, $M = 4$, each sensor has $N_i = 5$ antennae and transmission power $P_i = 2$, with observation noise level $\mathsf{SNR}_i = 9dB$. In our test, to take into account the impact of the channel parameters, for the above system set-up and any specific channel SNR we randomly generate 500 channel realizations. For each channel realization, two proposed algorithms are run, both of which start from one common random feasible solution. The progress of MI with respect to outer-loop iteration numbers are recorded. For one given iteration number, the average MI performance over all 500 channel realizations is presented in figure 6.2 and 6.3. For the implementation of 3-BCA algorithm, SDPT3 solver of CVX is chosen. The blue solid curves represent the average MI performance obtained by 3-block BCA algorithm with different numbers of iterations and the red dotted ones represent those obtained by cyclic $(L + 1)$ BCA algorithm. The black dotted curve represents the average MI obtained by random full-power-transmission solutions, which are actually the average MI performance for feasible solutions which make all power constraints active. From figures 6.2 and 6.3, we see that the optimized beamformers obtained by the proposed algorithms present significant MI improvement compared to nonoptimized beamformers.

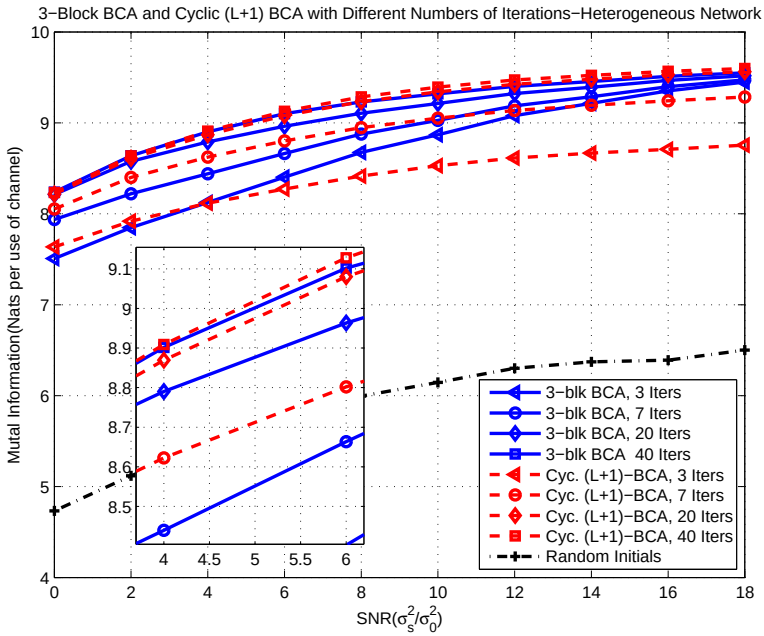

Figure 6.2: Heterogenous Test Case: 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Numbers of Iterations.

Usually 40 to 50 iteration loops are sufficient to make the two algorithms converge and the two algorithms finally converge to almost identical MI performance.

In figure 6.4 we test the special case of scalar source signal ($K = 1$), where $(L+1)$-block BCA algorithm has fully closed form solution, which is summarized in corollary 6.3.1. In this experiment, we have the system setup as follows $M = 4$, $N_1 = 3$, $N_2 = 4$, $N_3 = 5$, $P_1 = 1$, $P_2 = 2$, $P_3 = 3$ $\mathsf{SNR}_1 = 7dB$, $\mathsf{SNR}_2 = 8dB$ and $\mathsf{SNR}_3 = 9dB$. Similar results as in the multiple dimension source signal case have been obtained.

In figure 6.5 and 6.6, we check the impact of the random initials to the proposed algorithms. We use the same system setup as those in figure 6.2 and 6.3 respectively. Here the channel parameters are randomly chosen and fixed. 10 feasible solutions, each of which makes all the power constraints active(satisfied with equality) are randomly

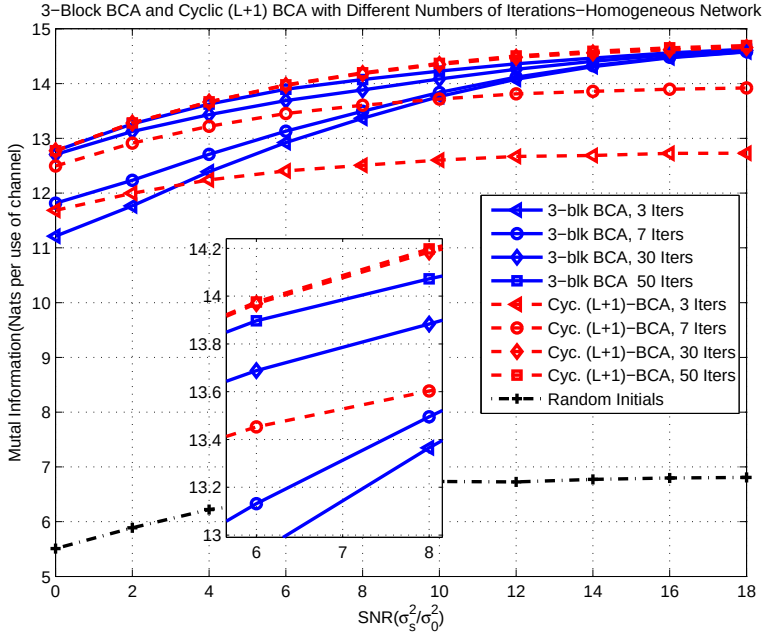

Figure 6.3: Homogeneous Test Case:3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Numbers of Iterations

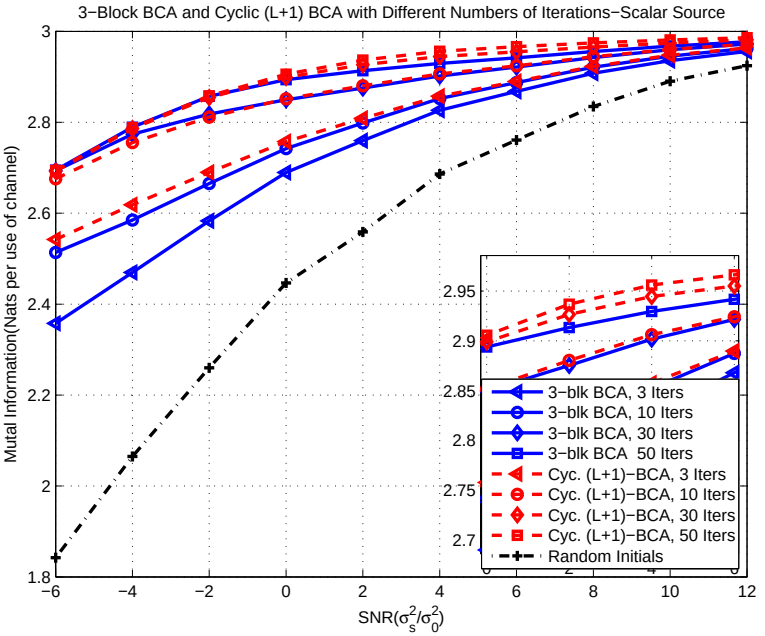

Figure 6.4: Scalar Source Signal Case: 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Numbers of Iterations.

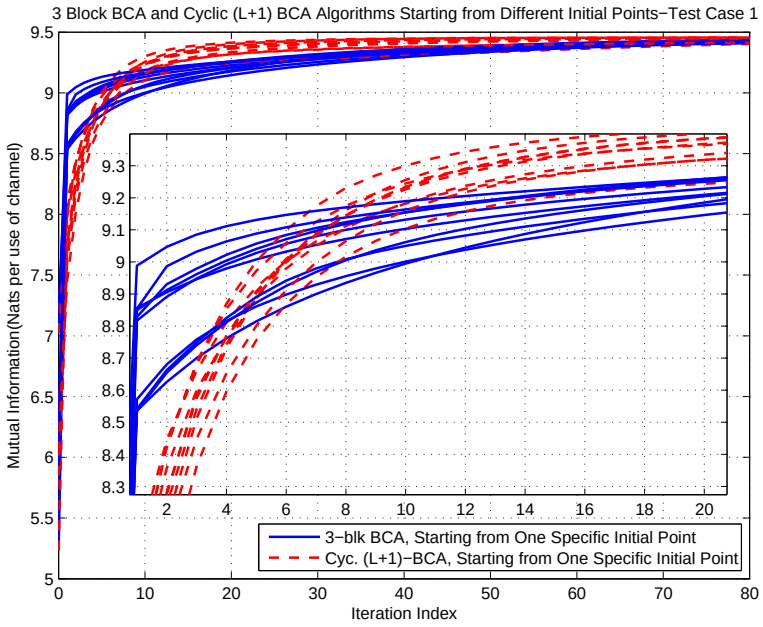

Figure 6.5: Heterogenous Test Case: Optimizing Generalized MI by 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Initial Points

generated. For each random initial point, we invoke the 3-block BCA and cyclic $(L+1)$-block BCA algorithms to optimize the beamformers. The MI progress of the proposed algorithms with different initials are illustrated in 6.5 and 6.6. It can be seen that both the 3-block BCA and cyclic $(L+1)$-block BCA are rather insensitive to selection of initial points. The two different algorithms with different initial points finally converge to almost identical value.

Last we test the complexity of the proposed algorithms. In Table 6.1 the average MATLAB running time for each out-loop is presented. Here we consider the homogeneous sensor network. Different values of K, L and N_i are tested, which result in different sizes of problem. SDPT3 solver of the CVX is used to implement 3-BCA algorithm. As shown in the table, the cyclic $(L+1)$ BCA algorithm is highly efficient, since each of its update step can be performed in an almost analytical way.

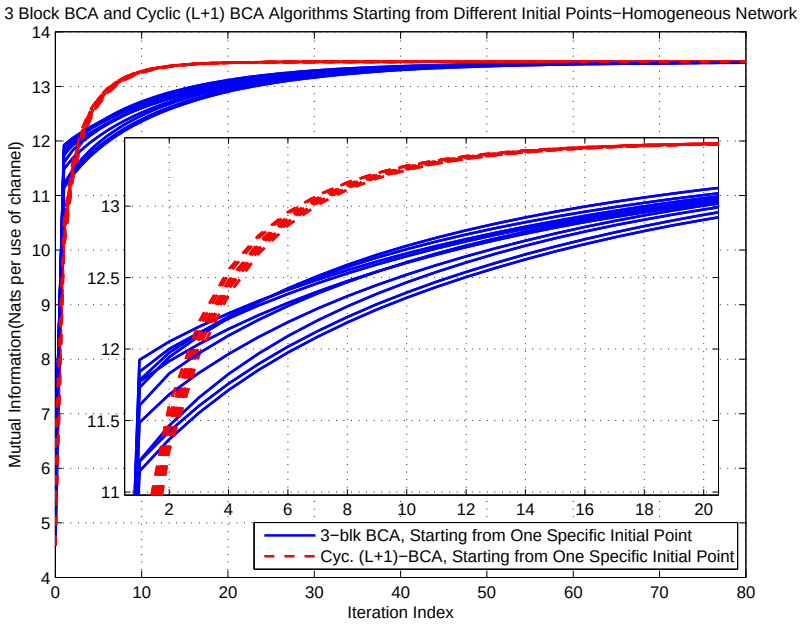

Figure 6.6: Homogeneous Test Case: Optimizing Generalized MI by 3-Block BCA Algorithm and Cyclic $(L+1)$-Block BCA Algorithm with Different Initial Points

Table 6.1: MATLAB Running Time Per Outer-Loop (in Sec.)

Dim. \ L	Algorithms	$L=5$	$L=10$	$L=20$
$K=1$	3-BCA	0.2748	0.4018	0.7627
$M=N_i=2$	Cyc. $(L+1)$	0.0125	0.0446	0.1792
$K=4$	3-BCA	0.5079	2.168	12.93
$M=N_i=4$	Cyc. $(L+1)$	0.0319	0.0999	0.3761
$K=8$	3-BCA	10.41	90.51	729.2
$M=N_i=8$	Cyc. $(L+1)$	0.0488	0.1310	0.4747

Notes: SDPT3 solver of CVX is chosen to implement 3-BCA.

6.5 Conclusion

In this chapter, we consider the linear beamforming design problem for a coherent MAC wireless sensor network to maximize the mutual information. As we have seen, the original problem is nonconvex and difficult. To solve this problem, we adopt the weighted minimum mean square error method and block coordinate ascent method to decompose the original difficult problem into subproblems and examine the solution to each subproblem, especially their closed form solution. The complexity and convergence of proposed algorithms are also discussed in details. Extensive numerical results are presented to verify and compare the behaviors of the proposed algorithms.

6.6 Appendix

6.6.1 Proof of Theorem 6.3.2

Proof. Since for each sub-problem, we solve an optimization problem with respect to a subset of variables with others being fixed, the objective value obtained by solving the current sub-problem cannot be smaller than previous one. Thus the entire MI sequence keeps increasing.

Under the positive definiteness assumption of $\mathbf{\Sigma_s}$, $\left(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\right) \succ 0$. Thus $\forall i \in \{1, \cdots, L\}$ we have

$$\|\mathbf{F}_i\|_F^2 \lambda_{\min}\left(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\right) \leq \mathsf{Tr}\{\mathbf{F}_i\left(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\right)\mathbf{F}_i^H\} \leq P_i, \tag{6.52}$$

where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of a Hermitian matrix. Since $\lambda_{\min}\big(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\big) > 0$, $\|\mathbf{F}_i\|_F^2$ is finite for all i. Thus the variable $\{\mathbf{F}_i\}_{i=1}^L$ is bounded. By Bolzano-Weierstrass theorem, there exits a subsequence $\{k_j\}_{j=1}^{\infty}$ such that $\{\mathbf{F}_i^{(k_j)}\}_{i=1}^L$ converges. Since $\mathbf{G}$ and $\mathbf{W}$ are updated by continuous functions of $\{\mathbf{F}_i\}_{i=1}^L$ in (6.24) and (6.23), $\big(\{\mathbf{F}_i^{(k_j)}\}_{i=1}^L, \mathbf{W}^{(k_j)}, \mathbf{G}^{(k_j)}\big)$ converges. Thus the existence of limit points in solution sequence has been proved.

We assume that $\big(\{\bar{\mathbf{F}}_i\}_{i=1}^L, \bar{\mathbf{W}}, \bar{\mathbf{G}}\big)$ is any limit point of $\big(\{\mathbf{F}_i^{(k)}\}_{i=1}^L, \mathbf{W}^{(k)}, \mathbf{G}^{(k)}\big)$. Then there exists a subsequence $\{k_j\}$ such that $\big(\{\mathbf{F}_i^{(k_j)}\}_{i=1}^L, \mathbf{W}^{(k_j)}, \mathbf{G}^{(k_j)}\big)$ $\stackrel{j\to\infty}{\longrightarrow} \big(\{\bar{\mathbf{F}}_i\}_{i=1}^L, \bar{\mathbf{W}}, \bar{\mathbf{G}}\big)$. Since $\{\mathbf{F}_i^{(k)}\}_{i=1}^L$ is bounded, by possibably restricting to a subsequence, we can assume that $\big(\{\mathbf{F}_i^{(k_j+1)}\}_{i=1}^L\big)$ converges to a limit $\big(\big\{\widehat{\mathbf{F}}_i\big\}_{i=1}^L\big)$.

Since for each j, $\{\mathbf{F}_i^{(k_j+1)}\}_{i=1}^L$ are feasible, i.e.

$$\mathsf{Tr}\{\mathbf{F}_i^{(k_j+1)}\big(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\big)\big(\mathbf{F}_i^{(k_j+1)}\big)^H)\} \le P_i, i \in \{1, \cdots, L\}. \tag{6.53}$$

By taking $j \to \infty$ in the above inequalities, we obtain

$$\mathsf{Tr}\{\widehat{\mathbf{F}}_i\big(\mathbf{\Sigma_s}+\mathbf{\Sigma}_i\big)\widehat{\mathbf{F}}_i^H)\} \le P_i, \;\; i \in \{1, \cdots, L\}. \tag{6.54}$$

So $\big\{\widehat{\mathbf{F}}_i\big\}_{i=1}^L$ are feasible.

For any feasible $\big\{\mathbf{F}_i\big\}_{i=1}^L$, we have

$$\mathsf{MI}\big(\{\mathbf{F}_i\}_{i=1}^L\big|\mathbf{W}^{(k_j)},\mathbf{G}^{(k_j)}\big) \le \mathsf{MI}\big(\{\mathbf{F}_i^{(k_j+1)}\}_{i=1}^L\big|\mathbf{W}^{(k_j)},\mathbf{G}^{(k_j)}\big). \tag{6.55}$$

Noticing that MI function is continuous and taking $j \to \infty$ in the above, we obtain

$$\mathsf{MI}\big(\{\mathbf{F}_i\}_{i=1}^L \big| \bar{\mathbf{W}}, \bar{\mathbf{G}}\big) \leq \mathsf{MI}\big(\{\widehat{\mathbf{F}}_i\}_{i=1}^L \big| \bar{\mathbf{W}}, \bar{\mathbf{G}}\big), \tag{6.56}$$

for any feasible $\{\mathbf{F}_i\}_{i=1}^L$.

Notice the $\{\mathbf{F}_i^{(k)}\}_{i=1}^L$ generated by algorithm 4 are feasible, by continuity of power constraint functions, $\{\bar{\mathbf{F}}_i\}_{i=1}^L$ are feasible. Thus we have

$$\mathsf{MI}\big(\{\bar{\mathbf{F}}_i\}_{i=1}^L \big| \bar{\mathbf{W}}, \bar{\mathbf{G}}\big) \leq \mathsf{MI}\big(\{\widehat{\mathbf{F}}_i\}_{i=1}^L \big| \bar{\mathbf{W}}, \bar{\mathbf{G}}\big). \tag{6.57}$$

At the same time, since the MI sequence is increasing and $\big(\{\bar{\mathbf{F}}_i\}_{i=1}^L, \bar{\mathbf{W}}, \bar{\mathbf{G}}\big)$ is a limit point of the solution sequence,

$$\mathsf{MI}\big(\{\bar{\mathbf{F}}_i\}_{i=1}^L \big| \bar{\mathbf{W}}, \bar{\mathbf{G}}\big) \geq \mathsf{MI}\big(\{\mathbf{F}_i^{(k)}\}_{i=1}^L \big| \bar{\mathbf{W}}, \bar{\mathbf{G}}\big), \tag{6.58}$$

for any integer k. Substitute k with k_j in (6.58), take limit $j \to \infty$ and combine it with (6.57), we have shown that $\{\bar{\mathbf{F}}_i\}_{i=1}^L$ is actually an optimal solution to the problem (P2) with parameters $\bar{\mathbf{W}}$ and $\bar{\mathbf{G}}$. So $\{\bar{\mathbf{F}}_i\}_{i=1}^L$ satisfy KKT conditions of (P2) with parameters $\bar{\mathbf{W}}$ and $\bar{\mathbf{G}}$, which are listed in (6.59) as follows

$$-\mathbf{H}_i^H \bar{\mathbf{G}} \bar{\mathbf{W}} \Big(\mathbf{I} - \bar{\mathbf{G}}^H \big(\sum_{i=1}^L \mathbf{H}_i \bar{\mathbf{F}}_i \big) \Big) \boldsymbol{\Sigma}_{\mathbf{s}} + \mathbf{H}_i^H \bar{\mathbf{G}} \bar{\mathbf{W}} \bar{\mathbf{G}}^H \mathbf{H}_i \bar{\mathbf{F}}_i \boldsymbol{\Sigma}_i$$

$$+ \lambda_i \bar{\mathbf{F}}_i \big(\boldsymbol{\Sigma}_{\mathbf{s}} + \boldsymbol{\Sigma}_i \big) = \mathbf{O}, \tag{6.59a}$$

$$\lambda_i \Big(\mathsf{Tr}\Big\{ \bar{\mathbf{F}}_i \big(\boldsymbol{\Sigma}_{\mathbf{s}} + \boldsymbol{\Sigma}_i \big) \bar{\mathbf{F}}_i^H \Big\} - P_i \Big) = 0, \tag{6.59b}$$

$$\mathsf{Tr}\Big\{ \bar{\mathbf{F}}_i \big(\boldsymbol{\Sigma}_{\mathbf{s}} + \boldsymbol{\Sigma}_i \big) \bar{\mathbf{F}}_i^H \Big\} \leq P_i, \tag{6.59c}$$

$$\lambda_i \geq 0. \tag{6.59d}$$

To simplify the following exposition, we introduce the following two notations:

$$\bar{\mathbf{H}} \triangleq \sum_{i=1}^{L} \mathbf{H}_i \bar{\mathbf{F}}_i; \tag{6.60a}$$

$$\bar{\mathbf{\Sigma}}_{\mathbf{n}} \triangleq \sigma_0^2 \mathbf{I} + \sum_{i=1}^{L} \mathbf{H}_i \bar{\mathbf{F}}_i \mathbf{\Sigma}_i \bar{\mathbf{F}}_i^H \mathbf{H}_i^H. \tag{6.60b}$$

According to the update step in algorithm 4, the limit points $\bar{\mathbf{W}}$ and $\bar{\mathbf{G}}$ have the relations with $\left\{\bar{\mathbf{F}}_i\right\}_{i=1}^{L}$ as follows.

$$\bar{\mathbf{G}} = \left[\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right]^{-1} \bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}, \tag{6.61a}$$

$$\bar{\mathbf{W}} = \bar{\mathbf{H}}^H \bar{\mathbf{\Sigma}}_{\mathbf{n}}^{-1} \bar{\mathbf{H}} + \mathbf{\Sigma}_{\mathbf{s}}^{-1}. \tag{6.61b}$$

Utilizing (6.61) we can prove two identities in (6.62) and (6.63) as follows

$$\begin{aligned}
&\bar{\mathbf{G}}\bar{\mathbf{W}}\left(\mathbf{I} - \bar{\mathbf{G}}^H \bar{\mathbf{H}}\right) \\
&= \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right)^{-1} \bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}} \left(\bar{\mathbf{H}}^H \bar{\mathbf{\Sigma}}_{\mathbf{n}}^{-1} \bar{\mathbf{H}} + \mathbf{\Sigma}_{\mathbf{s}}^{-1}\right) \left(\mathbf{I} - \mathbf{\Sigma}_{\mathbf{s}} \bar{\mathbf{H}}^H \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right)^{-1} \bar{\mathbf{H}}\right) \\
&= \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right)^{-1} \bar{\mathbf{H}} \\
&\quad + \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right)^{-1} \bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}} \Big[\underbrace{\bar{\mathbf{H}}^H \bar{\mathbf{\Sigma}}_{\mathbf{n}}^{-1} \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right) - \left(\bar{\mathbf{H}}^H \bar{\mathbf{\Sigma}}_{\mathbf{n}}^{-1} \bar{\mathbf{H}} + \mathbf{\Sigma}_{\mathbf{s}}^{-1}\right) \mathbf{\Sigma}_{\mathbf{s}} \bar{\mathbf{H}}^H}_{=\mathbf{O}} \Big] \\
&\qquad\qquad \cdot \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right)^{-1} \bar{\mathbf{H}} \\
&= \left(\bar{\mathbf{H}}\mathbf{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H + \bar{\mathbf{\Sigma}}_{\mathbf{n}}\right)^{-1} \bar{\mathbf{H}}
\end{aligned} \tag{6.62}$$

and

$$\begin{aligned}\bar{\mathbf{G}}\bar{\mathbf{W}}\bar{\mathbf{G}}^H &= \big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\big(\bar{\mathbf{H}}^H\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}^{-1}\bar{\mathbf{H}}+\boldsymbol{\Sigma}_{\mathbf{s}}^{-1}\big)\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H\big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\\ &= \big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\Big[\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}^{-1}\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H\Big]\big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\\ &= \big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\Big[\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}^{-1}\big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)\Big]\big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\\ &= \big(\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H+\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}\big)^{-1}\bar{\mathbf{H}}\boldsymbol{\Sigma}_{\mathbf{s}}\bar{\mathbf{H}}^H\bar{\boldsymbol{\Sigma}}_{\mathbf{n}}^{-1} \qquad (6.63)\end{aligned}$$

Substituting equations (6.62) and (6.63) into (6.59a), we can rewrite the first order KKT conditions associated with only $\{\bar{\mathbf{F}}_i\}_{i=1}^L$ in the following

$$\begin{aligned}\mathbf{H}_i^H\Big[\big(\sigma_0^2\mathbf{I}+\sum_{i=1}^L\mathbf{H}_i\bar{\mathbf{F}}_i\boldsymbol{\Sigma}_i\bar{\mathbf{F}}_i^H\mathbf{H}_i^H\big)+\big(\sum_{i=1}^L\mathbf{H}_i\bar{\mathbf{F}}_i\big)\boldsymbol{\Sigma}_{\mathbf{s}}\big(\sum_{i=1}^L\mathbf{H}_i\bar{\mathbf{F}}_i\big)^H\Big]^{-1}\big(\sum_{i=1}^L\mathbf{H}_i\bar{\mathbf{F}}_i\big)\boldsymbol{\Sigma}_{\mathbf{s}}\Big[\mathbf{I} \qquad (6.64)\\ -\big(\sum_{i=1}^L\mathbf{H}_i\bar{\mathbf{F}}_i\big)^H\big(\sigma_0^2\mathbf{I}+\sum_{i=1}^L\mathbf{H}_i\bar{\mathbf{F}}_i\boldsymbol{\Sigma}_i\bar{\mathbf{F}}_i^H\mathbf{H}_i^H\big)^{-1}\mathbf{H}_i\bar{\mathbf{F}}_i\boldsymbol{\Sigma}_i\Big]-\lambda_i\bar{\mathbf{F}}_i\big(\boldsymbol{\Sigma}_{\mathbf{s}}+\boldsymbol{\Sigma}_i\big)=\mathbf{O};\end{aligned}$$

To check the conditions of the original problem (P0), we need to determine the derivative of its Lagrangian function, or equivalently the derivative of MI with respect to $\{\mathbf{F}_i\}$. By defining

$$\mathbf{H}\triangleq\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i, \qquad (6.65)$$

the derivative of MI is calculated in (6.66) in the following with $C_1(\mathrm{d}\mathbf{F}_i)$ and $C_2(\mathrm{d}\mathbf{F}_i)$ being uninteresting terms involved $\mathrm{d}\mathbf{F}_i$ only and independent of $\mathrm{d}(\mathbf{F}_i^*)$.

$$
\begin{aligned}
\mathrm{d}(\mathsf{MI}) &= \mathsf{Tr}\Big\{\big(\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{H}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\big)^{-1}\mathrm{d}\big(\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{H}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\big)\Big\} \\
&= \mathsf{Tr}\Big\{\big(\mathbf{I}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{H}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\big)^{-1}\Big[\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathrm{d}(\mathbf{H}^H)\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{H}^H\mathrm{d}(\boldsymbol{\Sigma}_{\mathbf{n}}^{-1})\Big]\Big\} + C_1(\mathrm{d}\mathbf{F}_i) \\
&= \mathsf{Tr}\Big\{\mathbf{H}_i^H\big(\boldsymbol{\Sigma}_{\mathbf{n}}+\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\mathbf{H}^H\big)^{-1}\mathbf{H}\boldsymbol{\Sigma}_{\mathbf{s}}\Big[\mathbf{I}-\mathbf{H}^H\boldsymbol{\Sigma}_{\mathbf{n}}^{-1}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\Big]\mathrm{d}(\mathbf{F}_i)^H\Big\} + C_2(\mathrm{d}\mathbf{F}_i),
\end{aligned}
\tag{6.66a}
$$

$$
\begin{aligned}
\Rightarrow &\frac{\partial \mathsf{MI}}{\partial \mathbf{F}_i^*} = \\
&\mathbf{H}_i^H\Big[\big(\sigma_0^2\mathbf{I}+\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\big)+\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\boldsymbol{\Sigma}_{\mathbf{s}}\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)^H\Big]^{-1}\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)\boldsymbol{\Sigma}_{\mathbf{s}}\Big[\mathbf{I}- \\
&\qquad\big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\big)^H\big(\sigma_0^2\mathbf{I}+\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\big)^{-1}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\Big],\ i\in\{1,\cdots,L\}.
\end{aligned}
\tag{6.66b}
$$

By comparing the equations (6.64) with the derivative in (6.66b), it is easily to recognize that (6.64) is actually the first order KKT condition of problem (P0) optimizing MI. Together with equations (6.59b), (6.59c) and (6.59d), the KKT conditions of original problem have been proved to be satisfied by $\{\bar{\mathbf{F}}_i\}_{i=1}^{L}$. Thus the proof is complete. □

Chapter 7

Joint Transceiver Design towards SNR Maximization for Wireless Sensor Network

7.1 Introduction

In the last two chapters, we have discussed joint transceiver design problem in centralized wireless sensor networks with respect to metrics of mean square error(MSE) and mutual information(MI) respectively. MSE and MI are very standard and important criteria to evaluate the wireless sensor netowrks' performance from the perspectives of signal estimation deviation and network throughput. As we mentioned before, usually we have many different perspectives to assess the performance of wireless sensor network. In this chapter, besides of the two criteria discussed before, we turn to another

useful criterion—average signal to noise ratio(SNR) criterion. We know usually the observation obtained at the fusion center(FC) is corrupted by noise. When the source signal is a scalar and has finite alphabet, for example, it is finite-alphabet (e.g. QPSK or 16-QAM) or hybrid, then the forementioned criteria like MSE or MI are not suitable enough. In these cases, one solution is to consult to signal to noise ratio measure, which maximize the average power of the signals to that of noise. When the transmitted signal is one-dimension, the maximizing SNR criterion is equivalent to minimizing bit error ratio(BER) criterion, which is a standard signal detection criterion. So the method to jointly design transceivers to maximize average SNR motivates this chapter.

Nowadays the linear transceiver (beamforming) technique has been extensively studied and widely used in various kinds of communication systems. Taking advantage of linear filters at the transmitters and/or the receivers, the system performance can be significantly improved(e.g. spacial multiplexing gain). Linear transceivers provide sufficient flexibility for system design while maintain easiness for implementation. Under various system setup, the linear transceiver design in wireless sensor networks has raised many meaningful but also challenging problems, onto which enormous attention has been cast recently [48–53, 55, 71, 81, 82]. In [81] and [48], compression beamforming is considered for sensor networks, where the transmission channels from sensors to fusion center are assumed to be ideal(no fading and no noise). The model adopted in [81] and [48] is unrealistic since a typical wireless communication channel should be fading and noisy. Compression beamforming is also considered in [49], which makes a further step to jointly optimize precoders as well as the compression dimension assignment to each precoder. The research in [49] is performed under the assumption that noise covariance matrices at different sensors are identical (up to a scaling factor) and the all transmission channels are ideal. The beamforming problems under noisy channel

assumptions are considered in [50–53, 55, 71, 82]. [82] studies the simple case where each sensor-FC channel is a scalar fading channel with additive noise and all sensors share one total power constraint. Thus the beamforming problem in [82] boils down to a power allocation problem. [50] considers high dimension transmission channels under the assumptions that channel matrices are all square and nonsingular. As in [82], all sensors in [50] share transmission power and thus has one single power constraint. Note that the power sharing assumption in [82] and [50] is still stringent since in reality sensors are usually distributed in a sufficiently large area without wired connections between each other and powered separately by build-in batteries. [51] considers separate power constraint for each sensor under the assumption that all sensor-FC channels are square and nonsingular. Taking the additive channel noise and separate power constraint into account, [52] considers the special cases where all sensor-FC channels are scalars or identity matrices(nonfading). The most generic system models are considered in the recent work [53, 55, 71]. [53] and [71] consider the case that each sensor has separate power constraint and do not make any assumptions on beamformers' dimensions or channel matrices, i.e. each beamformer can be compressing, encoding or dimension-maintained and each channel matrix can be flat, slim or square (singular or nonsingular). All the above mentioned reference take mean square error (MSE) as performance metric. The beamforming problem aiming to maximize mutual information (MI) under the same generic system model is considered in [55].

The problem considered in this chapter is suitable for oversampling or cluster-based wireless sensor networks. Oversampling of the target is desirable since the observations of the sensors are usually corrupted by noise. More samples will lead to more accurate estimation/detection result at the fusion center. Multiple samples of the target are often obtainable in applications. For example when the target is slowly variant in time relative

to sensor's sampling procedure, sensor can perform sampling multiple times. Or each sensor can have multiple sets of sampling devices which can perform sampling within the same time slot. Furthermore it is interesting to note that mathematically, the over-sampling model is equivalent to cluster-based sensing problem discussed in [50]. For the cluster-based sensor network, each cluster is composed of a group of sensors which can communicate with others free of error and share transmission power(this happens in the case where sensors in the same cluster are located closely within a small nighborhood). Thus the head of each cluster collects observations from cluster members and then transmits all data to FC. In this chapter we adopt the general model used in [53,55,71], where the transmission channels are fading and noisy, each sensor or sensor cluster has separate transmission power constraint and no additional assumptions are posed on beamformers or channel matrices. Moreover we use the signal-to-noise ratio(SNR) as our performance metric, which is equivalent to symbol error rate(BER) metric for discrete signaling or channel capacity for Gaussian signaling. To solve the problem, we propose different iterative algorithms, with their convergence and complexity being examined and discussed.

The rest of this chapter is organized as follows. In section 7.2 we introduce our system model and formulate the beamforming design problem. Section 7.3 studies the optimal linear receiver. Section 7.4 proposes two 2-block coordinate ascent algorithms to solve our problem and section 7.5 introduces another framework which leads to low complexity algorithm in certain scenarios. Section 7.6 provides numerical results and section 7.7 concludes this chapter.

Notations: we use bold lowercase letters to denote complex vectors and bold capital letters to denote complex matrices. $\mathbf{0}$, $\mathbf{O}_{m\times n}$, and $\mathbf{I}_m$ are used to denote zero vectors,

zero matrices of dimension $m \times n$, and identity matrices of order m respectively. $\mathbf{A}^T$, $\mathbf{A}^*$ and $\mathbf{A}^H$ are used to denote transpose, conjugate and conjugate transpose(Hermitian transpose) respectively of an arbitrary complex matrix $\mathbf{A}$. $\mathsf{Tr}\{\cdot\}$ denotes the trace operation of a square matrix. $|\cdot|$ denotes the modulus of a complex scalar, and $\|\cdot\|_2$ denotes the l_2-norm of a complex vector. $\mathsf{vec}(\cdot)$ means vectorization operation of a matrix, which is performed by packing the columns of a matrix into a long one column. $\otimes$ denotes the Kronecker product. $\mathsf{Diag}\{\mathbf{A}_1, \cdots, \mathbf{A}_n\}$ denotes the block diagonal matrix with its i-th diagonal block being the square complex matrix $\mathbf{A}_i$, $i \in \{1, \cdots, n\}$. $\mathsf{Re}\{x\}$ and $\mathsf{Im}\{x\}$ denote the real and imaginary part of a complex value x respectively.

7.2 System Model

First we look at the system model for oversampling wireless sensor network, as shown in Fig.7.1. The system is composed of multiple sensors and one fusion center, all

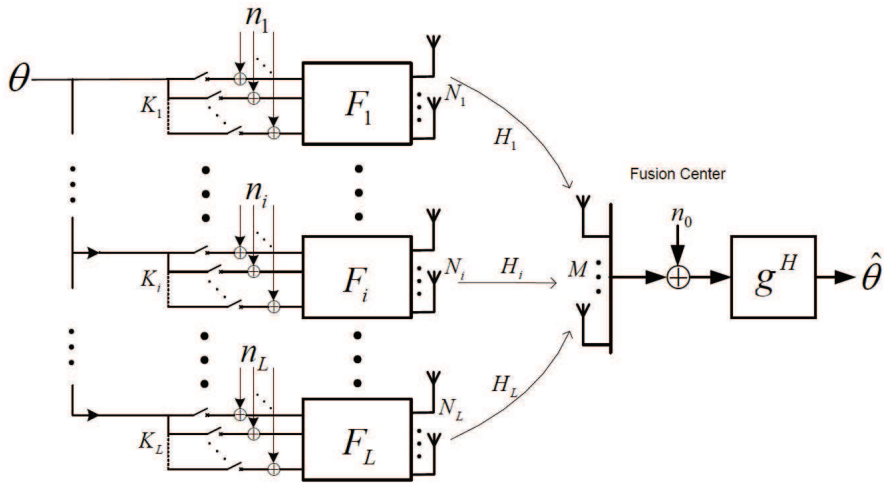

Figure 7.1: Models for Oversampling or Cluster-Based WSN

equipped with multiple antennas. N_i and M denote the number of antennas of the i-th

sensor and FC respectively. All sensors observe a common unknown parameter $\theta \in \mathbb{C}$. Without loss of generality, it is assumed that the target parameter θ has zero mean and unit covariance, i.e. $\mathsf{E}\{|\theta|^2\} = 1$. All sensors sample/oversample the target parameter and obtain noisy observations which are corrupted by thermal or environmental noise. The noisy observations are given as

$$\mathbf{x}_i = \mathbf{1}_{K_i}\theta + \mathbf{n}_i, \quad i = 1, \cdots, L, \tag{7.1}$$

where $\mathbf{1}_{K_i}$ is vector of dimension $K_i \times 1$ with all its entries being 1 and K_i denotes the number of samples. $\mathbf{n}_i$ denotes the observation noise with zero mean value and covariance matrices $\mathsf{E}\{\mathbf{n}_i\mathbf{n}_i^H\} = \mathbf{\Sigma}_i$. Here we make the mild assumption that $\mathbf{\Sigma}_i \succ 0$. The sensors in the network are sufficiently distributed in space thus it is reasonable to assume that the observations noise at different sensors are uncorrelated.

The noisy observation is beamformed by linear filter $\mathbf{F}_i \in \mathbb{C}^{N_i \times K_i}$ at each sensor before transmission. We denote the channel parameters from the i-th sensor to the fusion sensor as $\mathbf{H}_i \in \mathbb{C}^{M \times N_i}$. Here we assume that all the channel matrices are known at the fusion center, which can be obtained by standard channel estimation techniques. The received signal $\mathbf{r}$ at the fusion center reads:

$$\mathbf{r} = \sum_{i=1}^{L} \Big(\mathbf{H}_i\mathbf{F}_i\mathbf{x}_i\Big) + \mathbf{n}_0 \tag{7.2}$$

$$= \sum_{i=1}^{L} \Big(\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\Big)\theta + \Big(\sum_{i=1}^{L} \mathbf{H}_i\mathbf{F}_i\mathbf{n}_i + \mathbf{n}_0\Big), \tag{7.3}$$

where $\mathbf{n}_0$ is additive noise at the fusion center. Without loss of generality $\mathbf{n}_0$ has zero mean value and white covariance matrix, i.e. $\mathsf{E}\{\mathbf{n}_0\mathbf{n}_0^H\} = \sigma_0^2\mathbf{I}_M$. Since the fusion center is usually far away from the sensing field, $\mathbf{n}_0$ is uncorrelated with the observation noise

$\mathbf{n}_i$.

In practice, the wireless sensors are supported by build-in batteries or harvest energy from its surroundings, like polar energy or terrestiral heat. Thus due to different battery or environment conditions, each sensor has different transmission power constraint. The average transmission power for the i-th sensor is given as $\mathsf{E}\{\mathbf{F}_i(\mathbf{1}_{K_i}\theta + \mathbf{n}_i)(\mathbf{1}_{K_i}\theta + \mathbf{n}_i)^H\mathbf{F}_i^H\} = \mathsf{Tr}\big\{\mathbf{F}_i\big(\mathbf{1}_{K_i}\mathbf{1}_{K_i}^H + \mathbf{\Sigma}_i\big)\mathbf{F}_i^H\big\}$, which must be no greater than a power constraint P_i.

Note that the above signal model is also suitable for the cluster-based wireless sensor network, with each sensor in figure 7.1 being interpreted as a sensor cluster. In each cluster, the cluster head collects noisy samples from K_i cluster members as in (7.1), beamforms the observations and then transmits to FC. Each cluster has one separate power constraint and the noise between different clusters are uncorrelated.

At the fusion center, we utilize a linear receiver $\mathbf{g}$ to perform data fusion and obtain an estimate $\hat{\theta}$. The estimate obtained at the fusion center is given as

$$\hat{\theta}=\mathbf{g}^H\mathbf{r}=\Big(\mathbf{g}^H\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\Big)\theta+\mathbf{g}^H\Big(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{n}_i+\mathbf{n}_0\Big). \tag{7.4}$$

For the above sensor system, the merit of the recovered signal can be evaluated from different perspectives. One commonly used metric is mean square error(MSE), which is defined as $\mathsf{MSE} = \mathsf{E}\{|\theta - \hat{\theta}|^2\}$. When the target parameter θ comes from a continuous alphabet set, MSE serves as a standard metric for estimators' performance and the problem of jointly optimizing transceivers in wireless sensor network towards minimizing MSE is extensively discussed in existing literature [48–53, 71, 82]. One other important metric is average signal to noise ratio(SNR). When the target parameter

θ comes from a finite discrete alphabet set(e.g. M-PAM or M-QAM), the detection performance, which is measured by the symbol error probability(SER), usually has the form $\mathsf{SER} = c_1\mathsf{Q}\big(\sqrt{c_2\mathsf{SNR}}\big)$ with c_1 and c_2 being some positive constants. Another important problem is the mutual information $\mathsf{I}(\theta,\hat{\theta})$ between the target parameter and its estimate, which reads $\mathsf{I}(\theta,\hat{\theta}) = \frac{1}{2}\log_2(1+\mathsf{SNR})$. As seen above the transceiver design problems to optimize detection or mutual information reduce to SNR maximization problem. In (7.4) the linearly processed signal $\hat{\theta}$ is composed of signal component and noise component. The average SNR can be calculated as

$$\mathsf{SNR}\Big(\{\mathbf{F}\}_{i=1}^L,\mathbf{g}\Big) = \frac{\mathsf{E}\Big\{\Big|\Big(\mathbf{g}^H\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\Big)\theta\Big|^2\Big\}}{\mathsf{E}\Big\{\Big|\mathbf{g}^H\Big(\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{n}_i+\mathbf{n}_0\Big)\Big|^2\Big\}} \tag{7.5}$$

$$= \frac{\mathbf{g}^H\big[\sum_{i=1}^L\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\big]\big[\sum_{j=1}^L\mathbf{H}_j\mathbf{F}_j\mathbf{1}_{K_j}\big]^H\mathbf{g}}{\sigma_0^2\|\mathbf{g}\|_2^2+\sum_{i=1}^L\mathbf{g}^H\mathbf{H}_i\mathbf{F}_i\mathbf{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\mathbf{g}}, \tag{7.6}$$

where the assumptions of uncorrelated noise between the fusion center and different sensors have been invoked.

Thus the joint transceiver design problem maximizing SNR for oversampling or cluster-based wireless sensor network can be formulated as follows

$$(\mathsf{P0}): \max_{\{\mathbf{F}_i\}_{i=1}^L,\mathbf{g}\neq\mathbf{0}}.\ \mathsf{SNR}\Big(\{\mathbf{F}\}_{i=1}^L,\mathbf{g}\Big), \tag{7.7a}$$

$$\text{s.t. } \mathsf{Tr}\big\{\mathbf{F}_i\big(\mathbf{1}_{K_i}\mathbf{1}_{K_i}^H+\mathbf{\Sigma}_i\big)\mathbf{F}_i^H\big\}\leq P_i, i=1,\cdots,L. \tag{7.7b}$$

The problem $(\mathsf{P0})$ is a non-convex problem, which can be easily examined by checking the special case where all transceivers are scalars. In the following we consult to block coordinate ascent framework and propose iterative methods to solve the problem.

7.3 Optimal Linear Receiver

In this section, we obtain the optimal linear receiver $\mathbf{g}$ which leads to maximal SNR. The main result is as follows:

Theorem 7.3.1. For any predefined $\{\boldsymbol{F}_i\}_{i=1}^L$, SNR is maximized if and only if $\boldsymbol{g}^\star$ has the following form

$$\mathbf{g}^\star = \alpha\left(\sigma_0^2\mathbf{I}_M + \sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\right)^{-1}\left(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\right), \tag{7.8}$$

where α is arbitrary nonzero complex scalar. The maximal SNR is given as

$$\mathsf{SNR}^\star = \left\|\left(\sigma_0^2\mathbf{I}_M + \sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\right)^{-\frac{1}{2}}\left(\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\right)\right\|_2^2. \tag{7.9}$$

Proof. For simplicity we introduce the following notations

$$\mathbf{h} \triangleq \sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}; \tag{7.10a}$$

$$\mathbf{M} \triangleq \sigma_0^2\mathbf{I}_M + \sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\boldsymbol{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H. \tag{7.10b}$$

With all sensors' beamformers $\{\mathbf{F}_i\}_{i=1}^L$ given, the SNR maximization problem is the following optimization problem

$$\max_{\mathbf{g}\neq 0}\frac{\mathbf{g}^H\mathbf{h}\mathbf{h}^H\mathbf{g}}{\mathbf{g}^H\mathbf{M}\mathbf{g}}. \tag{7.11}$$

Since $\mathbf{M} \succ 0$, define $\tilde{\mathbf{g}} \triangleq \mathbf{M}^{\frac{1}{2}}\mathbf{g}$. The above problem becomes

$$\max_{\tilde{\mathbf{g}} \neq 0} . \frac{\tilde{\mathbf{g}}^H \mathbf{M}^{-\frac{1}{2}} \mathbf{h}\mathbf{h}^H \mathbf{M}^{-\frac{1}{2}} \tilde{\mathbf{g}}}{\tilde{\mathbf{g}}^H \tilde{\mathbf{g}}}. \tag{7.12}$$

From variational perspective, the maximal value of the above fractional is obtained if and only if $\tilde{\mathbf{g}}$ is aligned with eigen-vector corresponding to the maximal eigenvalue of the matrix $\mathbf{M}^{-\frac{1}{2}} \mathbf{h}\mathbf{h}^H \mathbf{M}^{-\frac{1}{2}}$ [92]. Notice that matrix $\mathbf{M}^{-\frac{1}{2}} \mathbf{h}\mathbf{h}^H \mathbf{M}^{-\frac{1}{2}}$ is rank-one and has only one positive eigenvalue whose eigen-vector is $\alpha \mathbf{M}^{-\frac{1}{2}} \mathbf{h}$, with α being any nonzero complex value. Thus the optimal solution of the above problem is $\tilde{\mathbf{g}}^\star = \alpha \mathbf{M}^{-\frac{1}{2}} \mathbf{h}$, from which (7.8) and (7.9) can be readily obtained. □

In practice the factor α can be chosen as 1 for convenience.

7.4 Jointly Optimizing Beamformers at Sensors

After obtaining the optimal linear receiver $\mathbf{g}$, we focus on optimizing precoders at the sensors' side in this section.

First by utilizing the identities $\mathsf{Tr}\{\mathbf{AB}\} = \mathsf{Tr}\{\mathbf{BA}\}$ and $\mathsf{Tr}\{\mathbf{ABCD}\} = \mathsf{vec}^H(\mathbf{D})[\mathbf{C}^T \otimes$

$\mathbf{A}\big]\mathsf{vec}\big(\mathbf{B}\big)$ [89], the numerator of SNR in (7.6) can be rewritten as follows

$$\begin{aligned}
&\mathbf{g}^H\Big[\sum_{i=1}^{L}\mathbf{H}_i\mathbf{F}_i\mathbf{1}_{K_i}\Big]\Big[\sum_{j=1}^{L}\mathbf{H}_j\mathbf{F}_j\mathbf{1}_{K_j}\Big]^H\mathbf{g}\\
&=\sum_{i,j=1}^{L}\mathsf{Tr}\Big\{\big(\mathbf{H}_j^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big)\mathbf{F}_i\big(\mathbf{1}_{K_i}\cdot\mathbf{1}_{K_j}^T\big)\mathbf{F}_j^H\Big\} \qquad (7.13)\\
&=\sum_{i,j=1}^{L}\mathsf{vec}^H\big(\mathbf{F}_j\big)\Big[\big(\mathbf{1}_{K_j}\cdot\mathbf{1}_{K_i}^T\big)\otimes\big(\mathbf{H}_j^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big)\Big]\mathsf{vec}\big(\mathbf{F}_i\big). \qquad (7.14)
\end{aligned}$$

Similarly the denominator of SNR can be written as

$$\begin{aligned}
&\sigma_0^2\|\mathbf{g}\|_2^2+\sum_{i=1}^{L}\mathbf{g}^H\mathbf{H}_i\mathbf{F}_i\mathbf{\Sigma}_i\mathbf{F}_i^H\mathbf{H}_i^H\mathbf{g}\\
&=\sum_{i=1}^{L}\mathsf{Tr}\Big\{\big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big)\mathbf{F}_i\mathbf{\Sigma}_i\mathbf{F}_i^H\Big\}+\sigma_0^2\|\mathbf{g}\|_2^2 \qquad (7.15)\\
&=\mathsf{vec}^H\big(\mathbf{F}_i\big)\Big[\mathbf{\Sigma}_i^*\otimes\big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big)\Big]\mathsf{vec}\big(\mathbf{F}_i\big)+\sigma_0^2\|\mathbf{g}\|_2^2, \qquad (7.16)
\end{aligned}$$

and the i-th power constraint is expressed as

$$\begin{aligned}
&\mathsf{Tr}\Big\{\mathbf{F}_i\Big(\mathbf{1}_{K_i}\cdot\mathbf{1}_{K_i}^T+\mathbf{\Sigma}_i\Big)\mathbf{F}_i^H\Big\} \qquad (7.17)\\
&=\mathsf{vec}^H\big(\mathbf{F}_i\big)\Big[\Big(\big(\mathbf{1}_{K_i}\cdot\mathbf{1}_{K_i}^T\big)+\mathbf{\Sigma}_i^*\Big)\otimes\mathbf{I}_{N_i}\Big]\mathsf{vec}\big(\mathbf{F}_i\big)\le P_i.
\end{aligned}$$

Here we introduce the following notations

$$\mathbf{f}_i \triangleq \mathsf{vec}\big(\mathbf{F}_i\big),\ \ i=1,\cdots,L; \tag{7.18a}$$

$$\mathbf{A}_{ij} \triangleq \Big[\big(\mathbf{1}_{K_i}\cdot\mathbf{1}_{K_j}^T\big)\otimes\big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_j\big)\Big],\ i,j=1,\cdots,L; \tag{7.18b}$$

$$\mathbf{B}_i \triangleq \Big[\mathbf{\Sigma}_i^*\otimes\big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big)\Big],\ \ i=1,\cdots,L; \tag{7.18c}$$

$$\mathbf{C}_i \triangleq \Big[\Big(\big(\mathbf{1}_{K_i}\cdot\mathbf{1}_{K_i}^T\big)+\mathbf{\Sigma}_i^*\Big)\otimes\mathbf{I}_{N_i}\Big],\ \ i=1,\cdots,L; \tag{7.18d}$$

$$c_0 \triangleq \sigma_0^2\|\mathbf{g}\|_2^2. \tag{7.18e}$$

and define the matrices $\mathbf{A}\triangleq[\mathbf{A}_{ij}]_{i,j=1}^L$, (i.e. the (i,j)-th elementary block of $\mathbf{A}$ is $\mathbf{A}_{ij}$), $\mathbf{B}\triangleq\mathsf{Diag}\{\mathbf{B}_1,\cdots,\mathbf{B}_L\}$ and $\mathbf{D}_i\triangleq\mathsf{Diag}\{\mathbf{O}_{\sum_{j=1}^{i-1}K_jN_j},\mathbf{C}_i,\mathbf{O}_{\sum_{j=i+1}^{L}K_jN_j}\}$ and pack all $\mathbf{f}_i$'s into one vector $\mathbf{f}\triangleq[\mathbf{f}_1^T,\cdots,\mathbf{f}_L^T]^T$. Then the problem of optimizing beamformers $\{\mathbf{F}_i\}_{i=1}^L$ with $\mathbf{g}$ given is reformulated as follows

$$(\mathsf{P1}):\max_{\mathbf{f}}.\ \frac{\mathbf{f}^H\mathbf{A}\mathbf{f}}{\mathbf{f}^H\mathbf{B}\mathbf{f}+c_0}, \tag{7.19a}$$

$$\mathsf{s.t.}\ \mathbf{f}^H\mathbf{D}_i\mathbf{f}\le P_i,\ \ i\in\{1,\cdots,L\}. \tag{7.19b}$$

In the following we discuss several methods solving the problem (P1).

7.4.1 Solving ($P1$) by Semidefinite Relaxation

In this subsection we solve problem (P1) with help of recent progress in semidefinite relaxation techniques.

First we rewrite the quadratic terms $\mathbf{f}^H\mathbf{A}\mathbf{f}$, $\mathbf{f}^H\mathbf{B}\mathbf{f}$ and $\mathbf{f}^H\mathbf{D}_i\mathbf{f}$ in (P1) into inner-product forms $\mathsf{Tr}\big\{\mathbf{A}\mathbf{X}\big\}$, $\mathsf{Tr}\big\{\mathbf{B}\mathbf{X}\big\}$ and $\mathsf{Tr}\big\{\mathbf{D}_i\mathbf{X}\big\}$ respectively by introducing the an

intermediate variable $\mathbf{X} = \mathbf{f}\mathbf{f}^H$. Omitting this rank-one constraint, a relaxation of (P1) is obtained as follows

$$(\mathsf{P2}): \max_{\mathbf{X}} . \frac{\mathsf{Tr}\{\mathbf{A}\mathbf{X}\}}{\mathsf{Tr}\{\mathbf{B}\mathbf{X}\} + c_0}, \tag{7.20a}$$

$$\text{s.t. } \mathsf{Tr}\{\mathbf{D}_i\mathbf{X}\} \leq P_i, \ \ i \in \{1, \cdots, L\}, \tag{7.20b}$$

$$\mathbf{X} \succcurlyeq 0. \tag{7.20c}$$

The fractional SDP objective in (P2) is still nonconvex. To solve it, we utilize Charnes-Cooper's approach, which was originally proposed in [83] and were widely adopted in fractional SDP optimization problems like [84, 85], to turn (P2) into the following SDP problem:

$$(\mathsf{P3}): \max_{\mathbf{Y},\nu} . \mathsf{Tr}\{\mathbf{A}\mathbf{Y}\}, \tag{7.21a}$$

$$\text{s.t. } \mathsf{Tr}\{\mathbf{B}\mathbf{Y}\} + c_0\nu = 1, \tag{7.21b}$$

$$\mathsf{Tr}\{\mathbf{D}_i\mathbf{Y}\} \leq P_i\nu, \ i \in \{1, \cdots, L\} \tag{7.21c}$$

$$\mathbf{Y} \succcurlyeq 0, \nu \geq 0. \tag{7.21d}$$

The equivalence between (P2) and (P3) is established by the following lemma.

Lemma 7.4.1. The problem $(P2)$ and $(P3)$ have equal optimal values. If $\boldsymbol{X}^\star$ solves $(P2)$, then $\left(\frac{\boldsymbol{X}^\star}{\mathsf{Tr}\{\boldsymbol{B}\boldsymbol{X}^\star + c_0\}}, \frac{1}{\mathsf{Tr}\{\boldsymbol{B}\boldsymbol{X}^\star + c_0\}}\right)$ is an optimal solution to $(P3)$. Conversely, if $(\boldsymbol{Y}^\star, \nu^\star)$ solves $(P3)$, then $\nu^\star > 0$ and $\boldsymbol{Y}^\star/\nu^\star$ solves $(P2)$.

Proof. See appendix 7.8.1. □

Since (P3) is SDP problem, it can be solved by standard numerical solvers like

CVX [69]. Remember that our goal is to solve problem (P1). If the optimal solution $\mathbf{Y}^\star$ to (P3) is rank-one, then the relaxation (P2) is tight to (P1) and the optimal solution of (P1) can be obtained by eigenvalue decomposition of $\mathbf{Y}^\star/\nu\star$. When the optimal solution $\mathbf{Y}^\star$ has rank larger than one, constructing a solution to (P1) from $(\mathbf{Y}^\star, \nu^\star)$ still needs to be addressed.

To introduce our first major conclusion we need the following lemma.

Lemma 7.4.2. The problem $(P3)$ and its dual are both solvable.

Proof. See appendix 7.8.2. □

For wireless sensor network with small number of sensors or sensor clusters, we have the following conclusion.

Theorem 7.4.1. If the wireless sensor network has no more than 3 sensors or sensor clusters, i.e. $L \leq 3$, then the relaxation $(P2)$ is tight with respect to $(P1)$. An optimal solution $(\boldsymbol{Y}^\star, \nu^\star)$ to $(P3)$ with $\boldsymbol{Y}^\star$ being rank-one can be constructed and solution to $(P1)$ can be obtained by eigenvalue-decomposing $\boldsymbol{Y}^\star/\nu^\star$.

Proof. The proof is inspired by theorem 3.2 of [87]. If $\mathbf{Y}^\star$ has rank one, nothing needs to be proved. Otherwise since the problem (P3) and its dual (D3) are both solvable by lemma 7.4.2, theorem 3.2 of [87] is valid to invoke. Define $r = \mathsf{rank}(\mathbf{Y}^\star)$ and perform the following procedure:

- **While** $\mathsf{rank}^2(\mathbf{Y}^\star) + \mathsf{rank}(\nu^\star) > L + 1$ **Do**

  Step-1: Perform a full rank decomposition $\mathbf{Y}^\star = \mathbf{V}\mathbf{V}^H$, where $\mathbf{V} \in \mathbb{C}^{(\sum_{i=1}^{L} K_i N_i) \times r}$;

Step-2: Find a nonzero pair $(\mathbf{\Delta}, \delta)$, where $\mathbf{\Delta}$ is a $r \times r$ Hermitian matrix and δ is real scalar, such that the following linear equations are satisfied

$$\mathsf{Tr}\{\mathbf{V}^H\mathbf{B}\mathbf{V}\mathbf{\Delta}\} + c_0\nu^\star\delta = 0; \tag{7.22}$$

$$\mathsf{Tr}\{\mathbf{V}^H\mathbf{D}_i\mathbf{V}\mathbf{\Delta}\} - P_i\nu^\star\delta = 0, i = 1, \cdots, L; \tag{7.23}$$

Step-3: Evaluate $\kappa = \max\Big(\big|\lambda_{min}(\mathbf{\Delta})\big|, \big|\lambda_{max}(\mathbf{\Delta})\big|, |\delta|\Big)$;

Step-4: Update $\mathbf{Y}^\star = \mathbf{V}\big(\mathbf{I}_{\sum_{i=1}^{L} K_i N_i} - \kappa^{-1}\mathbf{\Delta}\big)\mathbf{V}^H$, $\nu^\star = \nu^\star(1 - \kappa^{-1}\delta)$ and $r = \mathsf{rank}(\mathbf{Y}^\star)$;

- **End While**

In fact (P3) has two semidefinite variables $\mathbf{Y}$ and ν(note that ν is actually a nonnegative real scalar) and $L + 1$ constraints. As long as the condition $\mathsf{rank}^2(\mathbf{Y}^\star) + \mathsf{rank}(\nu)^\star > L + 1$ holds, nonzero solutions to (7.22) and (7.23) exist. Thus after each repetition a new optimal solution is constructed with $\mathsf{rank}(\mathbf{Y}^\star)$ being reduced by at least 1. Finally we obtain $\mathsf{rank}^2(\mathbf{Y}^\star) + \mathsf{rank}(\nu)^\star \leq L + 1$. Recall that $\nu^\star > 0$ by lemma 7.4.1, so $\mathsf{rank}(\nu)^\star = 1$ and we have $\mathsf{rank}^2(\mathbf{Y}^\star) \leq L \leq 3$. So $\mathsf{rank}(\mathbf{Y}^\star) = 1$ and the theorem is proved. □

In the above, we have seen that (P1) can be tackled by solving a SDP problem and then a finite number of linear equations when $L \leq 3$. However the assumption that $L \leq 3$ is still very stringent since in practice a sensor network can usually be composed of numerous sensors or clusters. A method to solve (P1) suitable for arbitrary L is still desirable. In the sequel, we proceed to discuss randomization method inspired by the

recent literature [88]. Before going into details, first we modify the problem (P3) a little bit. By changing the equality constraint (7.21b) into inequality, we have another SDP problem (P4) as follows

$$(\mathsf{P}4): \max_{\mathbf{Y},\nu} . \ \mathsf{Tr}\big\{\mathbf{AY}\big\}, \tag{7.24a}$$

$$\text{s.t. } \mathsf{Tr}\big\{\mathbf{BY}\big\} + c_0\nu \leq 1, \tag{7.24b}$$

$$\mathsf{Tr}\big\{\mathbf{D}_i\mathbf{Y}\big\} \leq P_i\nu, \ i \in \{1, \cdots, L\} \tag{7.24c}$$

$$\mathbf{Y} \succcurlyeq 0, \nu \geq 0. \tag{7.24d}$$

We assert that (P3) and (P4) are equivalent and for any solution $(\boldsymbol{Y}^\star, \nu^\star)$ to (P4), $\nu^\star$ must be positive. In fact since (P4) is a relaxation of (P3), $\mathsf{opt}(\mathsf{P}4) \geq \mathsf{opt}(\mathsf{P}3)$. Conversely, if $(\mathbf{Y}^\star, \nu^\star)$ is an optimal solution to (P4), then the constraint (7.24b) must indeed be active. Otherwise, $\mathbf{Y}^\star$ and $\nu^\star$ could be simultaneously inflated with a factor $\rho > 1$ such that $(\rho\mathbf{Y}^\star, \rho\nu^\star)$ satisfies all constraints with (7.24b) being active and gives an strictly larger objective, which contradicts the optimality of $(\mathbf{Y}^\star, \nu^\star)$. So $(\mathbf{Y}^\star, \nu^\star)$ is feasible for (P3) and thus $\mathsf{opt}(\mathsf{P}4) \leq \mathsf{opt}(\mathsf{P}3)$. Consequently (P3) and (P4) have equal optimal value. This means solution to either problem also solves the other one. Thus any solution $(\mathbf{Y}^\star, \nu^\star)$ to (P4) is also a solution to (P3) and by lemma 7.4.1, $\nu^\star > 0$.

Assuming that we have obtained an optimal solution $(\mathbf{Y}^\star, \nu^\star)$ to (P4), then we generate a sufficiently large number of independent complex random variables following the Gaussian distribution $\mathcal{CN}\big(\mathbf{0}, \mathbf{Y}^\star\big)$. The intuition of randomization lies in the following

observation. Consider the following stochastic optimization problem

$$(\mathsf{P}5): \max_{\mathbf{f},\nu} . \ \mathsf{E}_{\mathbf{f}\sim\mathcal{CN}(\mathbf{0},\mathbf{Y})}\{\mathbf{f}^H\mathbf{A}\mathbf{f}\}, \tag{7.25a}$$

$$\text{s.t. } \mathsf{E}_{\mathbf{f}\sim\mathcal{CN}(\mathbf{0},\mathbf{Y})}\{\mathbf{f}^H\mathbf{B}\mathbf{f}\} + c_0\nu \leq 1, \tag{7.25b}$$

$$\mathsf{E}_{\mathbf{f}\sim\mathcal{CN}(\mathbf{0},\mathbf{Y})}\{\mathbf{f}^H\mathbf{D}_i\mathbf{f}\} \leq P_i\nu, i=1,\cdots,L, \tag{7.25c}$$

$$\nu \geq 0. \tag{7.25d}$$

By utilizing the relation $\mathsf{E}_{\mathbf{f}\sim\mathcal{CN}(\mathbf{0},\mathbf{Y})}\{\mathbf{f}\mathbf{f}^H\} = \mathbf{Y}$, the stochastic problem $(\mathsf{P}5)$ is actually the SDP problem $(\mathsf{P}4)$. The random variable $\mathbf{f} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Y}^\star)$ solves the problem $(\mathsf{P}4)$ in expectation. Thus if we have sufficiently large number of samples, the "best" sample should solve the problem.

The "best" sample can be found as follows. Note that random samples are not always feasible for $(\mathsf{P}4)$. This issue can be addressed by the following rescaling procedure. For each sample $\tilde{\mathbf{f}}$, we define the scaling factor $\beta(\tilde{\mathbf{f}})$ as

$$\beta(\tilde{\mathbf{f}}) = \min_{i=1,\cdots,L} . \left\{1, \sqrt{\frac{1-c_0\nu^\star}{\tilde{\mathbf{f}}^H\mathbf{B}\tilde{\mathbf{f}}}}, \sqrt{\frac{P_i\nu^\star}{\tilde{\mathbf{f}}^H\mathbf{D}_i\tilde{\mathbf{f}}}}\right\}, \tag{7.26}$$

and rescale the sample $\tilde{\mathbf{f}}$ as

$$\bar{\mathbf{f}} = \frac{\beta(\tilde{\mathbf{f}})}{\sqrt{\nu^\star}}\tilde{\mathbf{f}}. \tag{7.27}$$

It is easy to check that the obtained $\bar{\mathbf{f}}$ is guaranteed to be feasible for $(\mathsf{P}4)$. Thus by performing the above rescaling procedure we can obtain ra large number of feasible samples to approximate the optimal solution $\mathbf{Y}^\star$. Then we choose the one giving maximal objective value as solution to the problem $(\mathsf{P}1)$. When the number of samples is

sufficiently large, the obtained best objective value of (P4) by rescaled random samples can be extremely close to true optimal value of (P4) so the randomization solution can be regarded as tight to the original problem (P1).

In retrospect the previous discussion, the motivation of transforming the problem (P3) into its equivalent (P4) now becomes clear. For implementation there is no chance that the randomly generated samples will satisfy the equality constraint (7.21b). At the same time the positivity of $\nu^\star$ guarantees that the rescaling in (7.27) can be performed.

Up to here, we have actually come out an alternative maximization method to solve the SNR optimization problem (P0) in (7.7). The algorithm starts from a random feasible point. In each iteration $\mathbf{g}$ is optimized in a closed form by theorem 7.3.1 with $\{\mathbf{F}_i\}_{i=1}^L$ being fixed and $\{\mathbf{F}_i\}_{i=1}^L$ are optimized by solving (P3) followed by randomization-rescaling or solving linear equations with $\mathbf{g}$ given.

This algorithm is summarized in algorithm 6 as follows.

7.4.2 Iteratively Solving ($P1$)

In the last subsection, we solve the problem (P1) with help of semidefinite relaxation by first solving its SDP relaxation and than construct the rank-one solution through solving linear equations or randomization method. In this subsection, we propose an alternative method which to solve (P1) in an iterative manner. First we have the following conclusion

Lemma 7.4.3. Matrix $\boldsymbol{A}$ in ($P1$) is rank-one. Specifically $\boldsymbol{A} = \boldsymbol{a}\boldsymbol{a}^H$ with the vector $\boldsymbol{a}$

Algorithm 6: 2-Block Coordinate Ascent Method to solve (P0) (SDR based)

Initialization: Randomly generate nonzero feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^L$ such that $\mathbf{g}^{(0)}$ obtained by theorem 7.3.1 is also nonzero; $j = 0$;
repeat
  Solve (P3) and obtain $(\mathbf{Y}^\star, \nu^\star)$;
  if $L \leq 3$ **then**
    Reducing rank of $\mathbf{Y}^\star$ as in theorem 7.4.1, obtain $\{\mathbf{F}_i^{(j+1)}\}_{i=1}^L$;
  else
    Generate sufficiently large number of samples following $\mathcal{CN}\big(\mathbf{0}, \mathbf{Y}^\star\big)$;
    Rescale each sample by (7.26) and (7.27);
    Select among all rescaled samples the one giving maximal SNR as $\{\mathbf{F}_i^{(j+1)}\}_{i=1}^L$;
  end
  Update $\mathbf{g}^{(j+1)}$ by theorem 7.3.1; $j++$;
until *increase of SNR is small enough or predefined number of iterations is reached*;

being given as

$$\boldsymbol{a} \triangleq \begin{bmatrix} \mathbf{1}_{K_1} \otimes \boldsymbol{H}_1^H \boldsymbol{g} \\ \vdots \\ \mathbf{1}_{K_L} \otimes \boldsymbol{H}_L^H \boldsymbol{g} \end{bmatrix}. \tag{7.28}$$

Proof. See appendix 7.8.3. □

Now looking at the fractional SDP objective of (P1) we have the following observation. For any given nonnegative real value γ, the SNR is no smaller than γ is equivalent to the fact

$$\mathbf{f}^H \mathbf{A} \mathbf{f} \geq \gamma \mathbf{f}^H \mathbf{B} \mathbf{f} + \gamma c_0. \tag{7.29}$$

In other words, if $\mathsf{opt}(\mathsf{P1}) \geq \gamma$, then there exits some $\mathbf{f}$ such that the inequality (7.29)

and all power constraints $\mathbf{f}^H\mathbf{D}_i\mathbf{f} \leq P_i$ for $i = 1, \cdots, L$ are simultaneously satisfied. If we define u as follows

$$u \triangleq \max_{i=1,\cdots,L} \left\{ \frac{\mathbf{f}^H\mathbf{D}_i\mathbf{f}}{P_i} \right\}, \tag{7.30}$$

then the fact that all power constraints are satisfied is equivalent to $u \leq 1$. Thus the statement $\mathsf{opt}(\mathsf{P1}) \geq \gamma$ holds if and only if the following optimization problem $(\mathsf{P6}_\gamma)$

$$(\mathsf{P6}_\gamma) : \min_{\mathbf{f}, u \geq 0} u \tag{7.31a}$$

$$\mathsf{s.t.}\ \mathbf{f}^H\mathbf{A}\mathbf{f} \geq \gamma\mathbf{f}^H\mathbf{B}\mathbf{f} + \gamma c_0, \tag{7.31b}$$

$$\frac{\mathbf{f}^H\mathbf{D}_i\mathbf{f}}{P_i} \leq u, i \in \{1, \cdots, L\}. \tag{7.31c}$$

has optimal value smaller than 1, i.e. $\mathsf{opt}(\mathsf{P6}_\gamma) \leq 1$.

Next we show that all constraints of problem $(\mathsf{P6}_\gamma)$ can be written in a second order cone form. The constraint (7.31b), utilizing the result of lemma 7.4.3, can be written as

$$\gamma\mathbf{f}^H\mathbf{B}\mathbf{f} + \gamma c_0 \leq |\mathbf{a}^H\mathbf{f}|^2 \tag{7.32}$$

Another key observation is that the optimal $\mathbf{f}^\star$ to $(\mathsf{P6}_\gamma)$ is phase invariant—$(\mathbf{f}^\star, u^\star)$ is optimal solution to $(\mathsf{P6}_\gamma)$ if and only if $(e^{\mathrm{j}\theta}\mathbf{f}^\star, u^\star)$ is optimal for any real value θ. So without loss of optimality we assume that $\mathbf{a}^H\mathbf{f} = v$ with v being a nonnegative real value. Thus the constraint (7.32) readily becomes the second order cone $\sqrt{\gamma}\left\| \left[\mathbf{f}^H\mathbf{B}^{\frac{1}{2}}, \sqrt{c_0}\right] \right\|_2 \leq v$. For the i-th power constraint in (7.31c), it can also be written in a second order cone form $P_i^{-1/2}\|\mathbf{D}_i^{\frac{1}{2}}\mathbf{f}\|_2 \leq u$. Thus the problem $(\mathsf{P6}_\gamma)$ can be equivalently written in a stan-

dard SOCP form:

$$(\mathsf{P7}_\gamma): \min_{\mathbf{f},u,v} . \ u, \tag{7.33a}$$

$$\text{s.t.} \left\| \begin{bmatrix} \sqrt{\gamma}\mathbf{B}^{\frac{1}{2}} & \mathbf{0} \\ \mathbf{0}^T & \sqrt{\gamma c_0} \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ 1 \end{bmatrix} \right\|_2 \leq v, \tag{7.33b}$$

$$\mathsf{Re}\{\mathbf{a}^H\mathbf{f}\} = v, \tag{7.33c}$$

$$\mathsf{Im}\{\mathbf{a}^H\mathbf{f}\} = 0, \tag{7.33d}$$

$$\left\| \sqrt{P_i^{-1}}\mathbf{D}_i^{\frac{1}{2}}\mathbf{f} \right\|_2 \leq u, \ i = 1, \cdots, L. \tag{7.33e}$$

Thus if we know that the $\mathsf{opt}(\mathsf{P1})$ lives in some interval, then $\mathsf{opt}(\mathsf{P1})$ can be determined by a bisection search—we set γ as middle point of the current search interval, if ($\mathsf{opt}(\mathsf{P7}_\gamma) \leq 1$), then $\mathsf{opt}(\mathsf{P1})$ can achieve higher value and γ is a lower bound of $\mathsf{opt}(\mathsf{P1})$. Otherwise γ upper-bounds $\mathsf{opt}(\mathsf{P1})$.

Now the remaining problem is to determine an interval containing $\mathsf{opt}(\mathsf{P1})$, from which the bisection search can start with. Since (P1) is maximization problem, any feasible solution gives a lower bound of $\mathsf{opt}(\mathsf{P1})$. The following lemma provides an upper bound of $\mathsf{opt}(\mathsf{P1})$.

Lemma 7.4.4. Optimal value of $(P1)$ has an upper bound as follows

$$\mathsf{opt}(\mathsf{P1}) \leq c_0^{-1} \left(\sum_{i=1}^{L} K_i \sqrt{\frac{P_i}{\lambda_{min}(\mathbf{C}_i)}} \left\| \mathbf{H}_i^H \mathbf{g} \right\|_2 \right)^2 . \tag{7.34}$$

Proof. See appendix 7.8.4. □

Thus we have obtained an alternative method to solve the original problem (P0),

which also falls in the 2-block coordinate ascent framework. This algorithm is summarized in algorithm 7.

Algorithm 7: 2-Block Coordinate Ascent Method to solve (P0) (SOCP based)

Initialization: Randomly generate nonzero feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^L$ such that $\mathbf{g}^{(0)}$ obtained by theorem 7.3.1 is also nonzero; $j = 0$;
repeat
  Obtain $bd_l = \mathsf{SNR}\big(\{\mathbf{F}_i^{(j)}\}_{i=1}^L, \mathbf{g}^{(j)}\big)$ and bd_u by (7.34);
  repeat
    Set $\gamma = (bd_u + bd_l)/2$; solve ($\mathsf{P7}_\gamma$);
    if $\mathsf{opt}(\mathsf{P7}_\gamma) \leq 1$ **then**
      $bd_l = \gamma$;
    else
      $bd_u = \gamma$;
    end
  until $(bd_u - bd_l)$ *is small enough*;
  $\gamma = bd_l$;
  Solve ($\mathsf{P7}_\gamma$) to update $\{\mathbf{F}_i^{(j+1)}\}_{i=1}^L$;
  Update $\mathbf{g}^{(j+1)}$ by theorem 7.3.1; $j++$;
until *increase of SNR is small enough or predefined number of iterations is reached*;

7.4.3 Convergence and Complexity

For the proposed algorithms in previous subsections, we have the following conclusion on convergence:

Theorem 7.4.2. The sequence of SNR obtained by algorithm 6 or 7 converges. Moreover the solution sequence generated by algorithm 6 or 7 has limit points and each limit point is a stationary point of problem $(P0)$.

Proof. See appendix 7.8.5. □

The complexity of the proposed algorithms is complicated since the whole network has too many factors (K_i's and N_i's) that impact the problem size. To simplify the analysis, we consider homogeneous sensor networks, where each sensor or sensor cluster has the same the number of sampling copies and antennas, i.e. $K_i = K$ and $N_i = N$ for all $i = 1, \cdots, L$.

According to [94], by use of primal-dual interior point method, the complexity to solve (P4) can be shown to be $\mathcal{O}\big(L^{4.5}K^{3.5}N^{3.5}\big)$. The complexity to update $\mathbf{g}$ by (7.8) lies in the matrix inversion operation, which has complexity $\mathcal{O}\big(M^3\big)$. In practice several thousands of random samples are sufficient to guarantee a satisfying tightness of the obtained rank-reduced solutions (usually within 10^{-4}) and the number of samples does not increase with size of wireless sensor network. So the complexity for each outer-layer iteration of the SDP based 2BCA algorithm is $\mathcal{O}\big(L^{4.5}K^{3.5}N^{3.5}+M^3\big)$.

By [93] the complexity for solving SOCP problem (P7$_\gamma$) is $\mathcal{O}\big(L^{3.5}K^3N^3\big)$. Note that each iteration of bisection search solves (P7$_\gamma$) for one time, so (P7$_\gamma$) is solved multiple times within one outer-layer iteration. Taking different channel conditions and levels of predefined precision into account, numerical results show that the number of times solving (P7$_\gamma$) varies between the narrow range $[25, 35]$ and thus can be considered as a constant. Thus the complexity of outer-layer SOCP based 2BCA algorithm is $\mathcal{O}\big(L^{3.5}K^3N^3+M^3\big)$.

7.5 Multiple Block Framework to Maximize SNR

In the previous sections, the proposed algorithms are both 2-block coordinate ascent methods where all the beamformers' sensors are jointly updated. One problem for these algorithms is that the complexity of solving the associated SDP or SOCP problem grows intensively with the increase of the size of the wireless sensor network. Instead of jointly optimizing all beamformers, we can alternatively focus on just one sensor's beamformer each time. This actually results in a multiple-block coordinate ascent method. As it will be seen by complexity analysis and numerical results, this method can often significantly decrease complexity. Furthermore in some special but important case each block can be updated in a closed form, which does not depend on any numerical solver and has extremely low complexity.

Now we consider the problem of optimizing the i-th beamformer $\mathbf{F}_i$ with $\mathbf{g}$ and $\{\mathbf{F}_j\}_{j\neq i}$ being fixed. By introducing the following notations

$$\mathbf{q}_i \triangleq \sum_{j\neq i} \mathbf{A}_{ij}\mathbf{f}_j;\ c_i \triangleq \sum_{j,k\neq i} \mathbf{f}_j^H \mathbf{A}_{jk}\mathbf{f}_k; \tag{7.35a}$$

$$d_i \triangleq \sigma_0^2\|\mathbf{g}\|_2^2 + \sum_{j\neq i} \mathbf{f}_j \mathbf{B}_j \mathbf{f}_j, \tag{7.35b}$$

this problem is formulated as follows

$$(\mathsf{P1}^i): \max_{\mathbf{f}_i} . \ \frac{\mathbf{f}_i^H \mathbf{A}_{ii}\mathbf{f}_i + 2\mathsf{Re}\{\mathbf{q}_i^H\mathbf{f}_i\} + c_i}{\mathbf{f}_i^H \mathbf{B}_i \mathbf{f}_i + d_i}, \tag{7.36a}$$

$$\mathsf{s.t.}\ \mathbf{f}_i^H \mathbf{C}_i \mathbf{f}_i \leq P_i. \tag{7.36b}$$

7.5.1 One-Shot SDR-Rank-Reduction Method

First we introduce a one-shot method to solve ($\mathsf{P}1^i$), which performs semidefinite programming and rank-one matrix decomposition in tandem. This method is discussed in recent work [84] and [85].

By use of Charnes-Cooper's transformation and rank-one relaxation we turn ($\mathsf{P}1^i$) into the following relaxed version

$$(\mathsf{P}7^i): \max_{\mathbf{Z},\eta} . \ \mathsf{Tr}\big\{\mathbf{Q}_1\mathbf{Z}\big\}, \tag{7.37a}$$

$$\text{s.t. } \mathsf{Tr}\big\{\mathbf{Q}_2\mathbf{Z}\big\} = 1, \tag{7.37b}$$

$$\mathsf{Tr}\big\{\mathbf{Q}_3\mathbf{Z}\big\} \leq P_i\eta, \tag{7.37c}$$

$$\mathsf{Tr}\big\{\mathbf{Q}_4\mathbf{Z}\big\} = \eta, \tag{7.37d}$$

$$\mathbf{Z} \succcurlyeq 0, \eta \geq 0. \tag{7.37e}$$

with parameter matrices being defined as

$$\mathbf{Q}_1 \triangleq \begin{bmatrix} \mathbf{A}_{ii} & \mathbf{q}_i \\ \mathbf{q}_i^H & c_i \end{bmatrix}, \mathbf{Q}_2 \triangleq \begin{bmatrix} \mathbf{B}_i & \mathbf{0} \\ \mathbf{0}^T & d_i \end{bmatrix}, \tag{7.38a}$$

$$\mathbf{Q}_3 \triangleq \begin{bmatrix} \mathbf{C}_i & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \mathbf{Q}_4 \triangleq \begin{bmatrix} \mathbf{O} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}. \tag{7.38b}$$

Solving the SDP ($\mathsf{P}7^i$) we obtain an solution $(\mathbf{Z}^\star, \eta^\star)$. If the $\mathbf{Z}^\star$ is rank one, i.e. $\frac{\mathbf{Z}^\star}{\eta^\star} = \mathbf{z}^\star\mathbf{z}^{\star H}$ with $\mathbf{z}^\star \triangleq [\mathbf{z}_1^T, z_2]^T$, then $\mathbf{z}_1^\star/z_2$ is an solution to ($\mathsf{P}1^i$) and the relaxation ($\mathsf{P}7^i$) is actually tight with respect to ($\mathsf{P}1^i$). Actually the rank-one solution $\mathbf{Z}^\star$ always exits due to the recent matrix decomposition result in [86]. In fact if $\mathbf{Z}^\star$ has rank larger than

one, by help of theorem 2.2 in [86], we can obtain a vector $\mathbf{z}$ such that the equations $\mathsf{Tr}\{(\mathbf{Q}_1 - \mathsf{opt}(\mathsf{P7}^i)\mathbf{Q}_2)\mathbf{z}\mathbf{z}\} = 0$, $\mathsf{Tr}\{\mathbf{Q}_j\mathbf{z}\mathbf{z}^H\} = \mathsf{Tr}\{\mathbf{Q}_j\mathbf{Z}^\star\}$ for $j = 3, 4$. This means $(\mathbf{z}\mathbf{z}, \eta^\star)$ is rank-one optimal solution to $(\mathsf{P7}^i)$ and thus $(\mathsf{P1}^i)$ can be solved.

7.5.2 Iterative Method

Besides the above one-shot method, here we propose an alternative iterative method to solve $(\mathsf{P1}^i)$. As we will see this iterative method can give birth to extremely efficient solution to $(\mathsf{P1}^i)$ in specific circumstance.

For any given positive real value α, the fact that the SNR objective in $(\mathsf{P1}^i)$ is no smaller than α equivalently reads

$$\mathbf{f}_i^H\big[\alpha\mathbf{B}_i - \mathbf{A}_{ii}\big]\mathbf{f}_i - 2\mathsf{Re}\big\{\mathbf{q}_i^H\mathbf{f}_i\big\} + (\alpha d_i - c_i) \leq 0. \tag{7.39}$$

This immediately implies that if the following problem with

$$(\mathsf{P8}_\alpha^i): \min_{\mathbf{f}_i} . \ \mathbf{f}_i^H\big[\alpha\mathbf{B}_i - \mathbf{A}_{ii}\big]\mathbf{f}_i - 2\mathsf{Re}\big\{\mathbf{q}_i^H\mathbf{f}_i\big\} + (\alpha d_i - c_i), \tag{7.40a}$$

$$\mathsf{s.t.}\ \mathbf{f}_i^H\mathbf{C}_i\mathbf{f}_i \leq P_i. \tag{7.40b}$$

with α given has a nonnegative optimal value then $\mathsf{opt}(\mathsf{P1}^i) \geq \alpha$. Otherwise α can serve as an upper bound of $\mathsf{opt}(\mathsf{P1}^i)$. Thus we can perform a bisection search to solve $(\mathsf{P1}^i)$. Now the problem reduces to how to solve the problem $(\mathsf{P8}_\alpha^i)$? Note that the quadratic matrix $\big[\alpha\mathbf{B}_i - \mathbf{A}_{ii}\big]$ can be negative semidefinite or indefinite and thus $(\mathsf{P8}_\alpha^i)$ is possibly nonconvex. The following theorem convinces us that $(\mathsf{P8}_\alpha^i)$ can always be solved regardless of the convexity of its objective.

Theorem 7.5.1. If the i-th sensor or sensor cluster has multiple copies of observation or is equipped with multiple antenna, i.e. $K_i \geq 2$ or $N_i \geq 2$, the problem $(P8^i_\alpha)$ can be solved.

Proof. See appendix 7.8.6. □

Although theorem 7.5.1 shows that the problem $(\mathsf{P}8^i_\alpha)$ can be solved by SDR and thus the iterative method to solve $(\mathsf{P}1^i)$ works, it is generally less efficient than the one-shot method discussed above. Since the former performs semidefinite programming and rank-one reductions multiple times while the latter for just once. However in the circumstance where $K_i = 1$, the following theorem indicates that $(\mathsf{P}8^i_\alpha)$ has fully closed form solution and consequently the iterative method can become extremely efficient.

Theorem 7.5.2. When $K_i = 1$ the solution $\boldsymbol{f}_i^\star$ to $(P8^i_\alpha)$ is

$$\boldsymbol{f}_i^\star = \begin{cases} \frac{\beta_i^* \mathbf{H}_i^H \mathbf{g}}{\|\mathbf{H}_i^H \mathbf{g}\|_2^2(\alpha\sigma_i^2 - 1)}, & \text{if } \alpha\sigma_i^2 > 1,\ \frac{|\beta_i^*|}{\|\mathbf{H}_i^H \mathbf{g}\|_2(\alpha\sigma_i^2 - 1)} \leq \sqrt{\bar{P}_i}; \\ \frac{\sqrt{\bar{P}_i}\beta_i^* \mathbf{H}_i^H \mathbf{g}}{|\beta_i| \|\mathbf{H}_i^H \mathbf{g}\|_2}, & \text{otherwise,} \end{cases}$$

with β_i and $\bar{P}_i$ being defined by equation (7.69) in the appendix.

Proof. See appendix 7.8.7. □

Note that the closed form solution in theorem 7.5.2 neither requires matrix decomposition or solving linear equations(matrix inversion) nor depends on numerical solver. Thus iteratively solving $(\mathsf{P}1^i)$ is easy for implementation and has very low computation cost. Comparatively the one-shot method to solve $(\mathsf{P}1^i)$ depends on numerical

solvers(like CVX) which are in fact iterative-based(interior point method) solvers and each iteration performs matrix decomposition and solving linear equations.

To start the bisection search, the latest SNR can serve as a lower bound for $\mathsf{opt}(\mathsf{P}1^i)$. By equation (7.56) it can be shown that an upper bound for $\mathsf{opt}(\mathsf{P}1^i)$ is given as

$$d_i^{-1}\Big(\|\mathbf{H}_i^H\mathbf{g}\|_2^2 P_i \lambda_{min}^{-1}\big(\mathbf{C}_i\big)+2\|\mathbf{q}_i\|_2 P_i^{\frac{1}{2}}\lambda_{min}^{-\frac{1}{2}}\big(\mathbf{C}_i\big)+c_i\Big). \tag{7.41}$$

Note that the above upper bound can be much tighter than the one given in (7.34) since it utilizes the knowledge of $\{\mathbf{F}_j\}_{j\neq i}$.

From the above discussion, we can utilize multiple block coordinate ascent method to solve the original SNR problem ($\mathsf{P}0$). For each update we optimize one separate sensor's precoder or the FC receiver. For the i-th sensor, if $K_i > 1$, its beamformer can be updated by the one-shot SDR-rank-reduction method. Otherwise theorem 7.5.2 can be invoked. This is summarized in algorithm 8.

Although the multiple BCA method generates monotonically increasing SNR sequence, it is hard to prove that the limit points of its solution sequence guarantee to converge to stationary points of ($\mathsf{P}0$). Numerical results in section 7.6 show that multiple BCA algorithm usually has a very satisfying convergence behaviors.

By primal-dual inter point method [94], the complexity of each outer-layer iteration of multiple BCA for homogeneous wireless sensor network is $\mathcal{O}\big(LK^{3.5}N^{3.5}+LM^3\big)$. Particularly for homogeneous network with $K=1$, the complexity becomes $\mathcal{O}\big(LM^3\big)$ with the help of theorem 7.5.2.

Algorithm 8: Multiple-Block Coordinate Ascent Method to solve ($\mathsf{P}0$)

Initialization: Randomly generate nonzero feasible $\{\mathbf{F}_i^{(0)}\}_{i=1}^L$ such that $\mathbf{g}^{(0)}$ obtained by (7.8) is nonzero;
repeat
  for $i = 1, \cdots, L$; **do**
    if $K_i > 1$ **then**
      Solve ($\mathsf{P}7^i$) then perform rank-reduction by theorem 2.2 in [86]; update $\mathbf{F}_i$;
    else
      Set bd_l as current SNR; obtain bd_u by (7.41);
      repeat
        Set $\alpha = (bd_u + bd_l)/2$;
        Solve ($\mathsf{P}8_\alpha^i$) by theorem (7.5.2);
        if $\mathsf{opt}(\mathsf{P}8_\alpha^i) \leq 0$ **then**
          $bd_l = \alpha$;
        else
          $bd_u = \alpha$;
        end
      until $(bd_u - bd_l)$ *is small enough*;
      $\alpha = bd_l$;
      solve ($\mathsf{P}8_\alpha^i$) by theorem 7.5.2; update $\mathbf{F}_i$;
    end
    Update $\mathbf{g}$ by theorem 7.3.1 ;
  end
until *increase of SNR is small enough or predefined number of iterations is reached*;

7.6 Numerical Results

In this section, numerical results are presented to testify the proposed algorithms' performance. In our experiments, the observation noise at each sensor is colored, which has a covariance

$$\mathbf{\Sigma}_i = \sigma_i^2 \mathbf{\Sigma}_0, \quad i \in \{1, \cdots, L\}, \tag{7.42}$$

where the $K_i \times K_i$ matrix $\mathbf{\Sigma}_0$ has the Toeplitz structure

$$\mathbf{\Sigma}_0 = \begin{bmatrix} 1 & \rho & \ddots & \rho^{K-1} \\ \rho & 1 & \ddots & \ddots \\ \ddots & \ddots & \ddots & \rho \\ \rho^{K-1} & \ddots & \rho & 1 \end{bmatrix}. \tag{7.43}$$

The parameter ρ is set to 0.5 for all sensors in the following experiments. Here we define the observation signal to noise ratio at the i-th sensor as $\mathsf{SNR}_i \triangleq \sigma_i^{-2}$ and the channel signal to noise ratio as $\mathsf{SNR} \triangleq \sigma_0^{-2}$.

In figure 7.2 and 7.3 the average SNR obtained at the FC are plotted. It is assumed that the sensor network has 5 sensors and FC has 4 antennas. We set $N_1 = 3, N_2 = 4, N_3 = 5, N_4 = 4, N_5 = 5$, $K_1 = 3, K_2 = 4, K_3 = 5, K_4 = 6, K_5 = 6$ and $P_1 = 0.2, P_2 = 0.2, P_3 = 0.3, P_4 = 0.2, P_5 = 0.3$. For each fixed channel SNR level, 50 random channel realizations are generated with each element of channel matrix follows circularly symmetric complex Gaussian distribution with zero mean and covariance 2. With channel SNR and channel matrices given, the proposed algorithms are performed

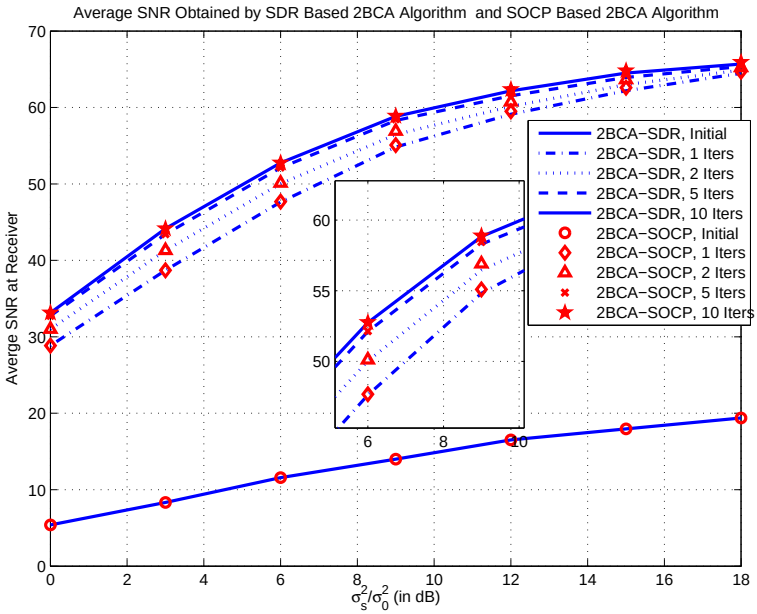

Figure 7.2: Average SNR Obtained by SDR Based 2BCA Algorithm and SOCP Based 2BCA Algorithm

starting from one common random initial. The obtained average SNR is plotted in figure 7.2 and 7.3. The obtained average SNR of SDR based and SOCP based 2BCA algorithms are plotted in figure 7.2 with respect to different outer iterations. The curve associated with random initials actually represents the performance of random feasible linear transmitters. From figure 7.2, optimized SNR converges in 10 outer-iterations on average. These two algorithms have identical average convergence performance, this will also be verified by figure 7.4. The average SNR performance obtained by multiple BCA algorithm is presented in figure 7.3, where SDR based 2BCA algorithm serves as a benchmark. Multiple BCA algorithm presents identical average SNR performance with the other 2 block algorithms.

In figure 7.4 and 7.5, the impact of different initial points to the algorithms are examined. The system setup is identical with the experiment in figure 7.2 and 7.3. We set

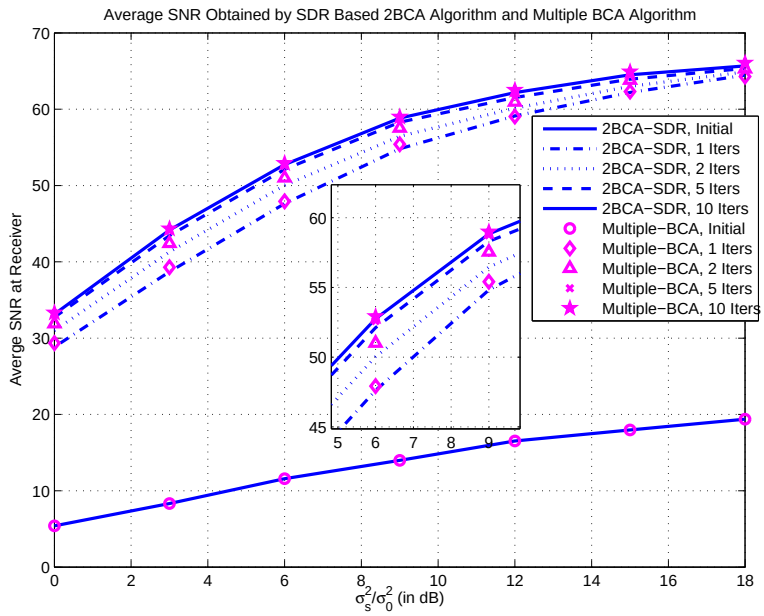

Figure 7.3: Average SNR Obtained by SDP Based 2BCA Algorithm and Multiple BCA Algorithm

the channel SNR as 2dB and fix the channel matrices with one specific random realization. The three proposed algorithms are started from 10 different random initials and each SNR itinerary with respect to number of outer-layer iterations is plotted in figure 7.4 and 7.5, where the itineraries of SDR based 2BCA algorithm serve as benchmarks. From figure 7.4 it can be seen that the two 2BCA algorithms have almost identical SNR itineraries. Comparatively, multiple BCA algorithm's itineraries are usually very different but finally it will converge to identical value. Figures 7.4 and 7.5 reflect the fact that: the proposed three algorithms are initial-insensitive; they finally converge to identical SNR value; and usually 30 iterations are sufficient for these proposed algorithms to converge.

Next we present numerical results for complexity. Still we take homogeneous wireless sensor network as example. N and M denote the number of antennas for each

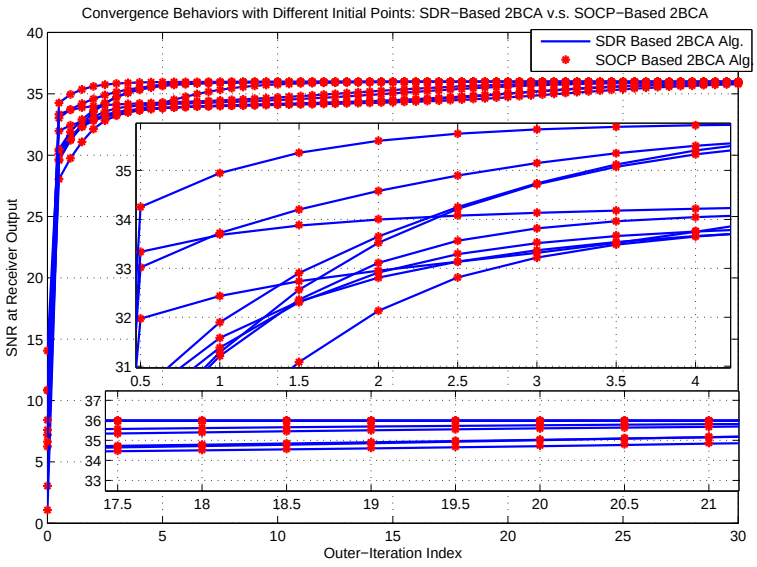

Figure 7.4: Convergence with Different Initials: SDR Based 2BCA Algorithm v.s. SOCP Based 2BCA Algorithm

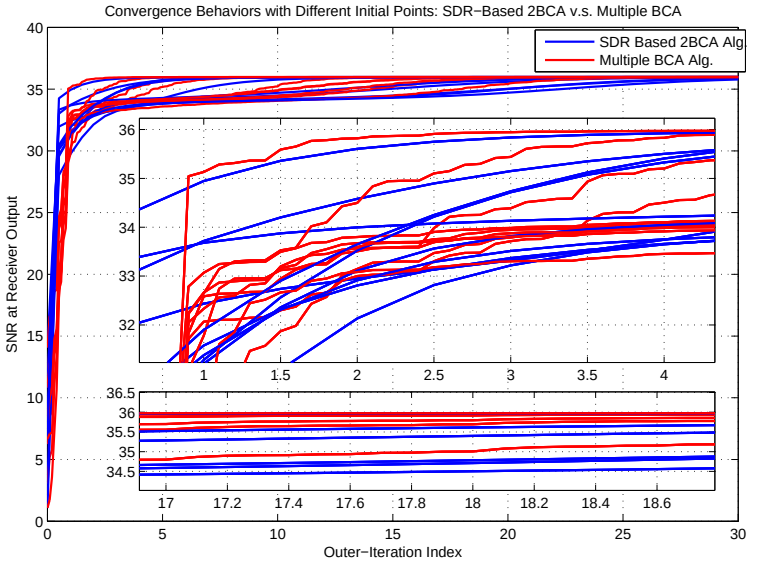

Figure 7.5: Convergence with Different Initials: SDR Based 2BCA Algorithm v.s. Multiple BCA Algorithm

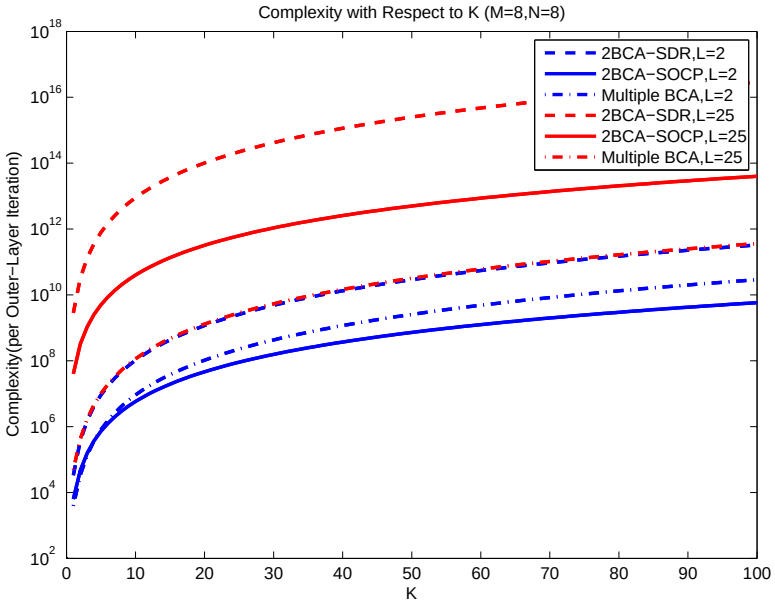

Figure 7.6: Complexity of Algorithms with Respect to K

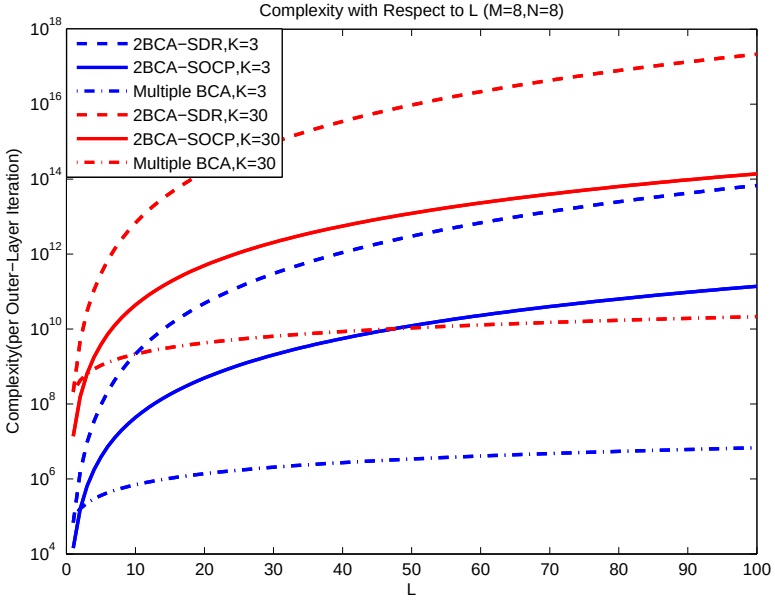

Figure 7.7: Complexity of Algorithms with Respect to L

sensor or sensor cluster and FC respectively and take modest values within several tens. Comparatively the number of sensors or sensor clusters can be large, and one cluster can have numerous sensors. So we focus on the impact of L and K on the complexity. Figure 7.6 and 7.7 represent the complexity for each outer-layer iteration for proposed algorithms with respect to K and L respectively. Generally SDR based 2BCA algorithm has higher complexity than the two others. The SOCP based algorithm has lowest complexity for large K with small L and multiple BCA algorithm has the lowest complexity for large L.

In the following the average execution time of proposed algorithms using MATLAB with the standard toolbox CVX v2.1 on the same computer are presented in table 7.1 and 7.2. The multiple BCA algorithm requires much lower time for networks with large L and SOCP based 2BCA algorithm is more efficient for large K and small L. Although the complexity of SDR-based 2BCA algorithm increases drastically with the increase of K, N and L in general, it can still be useful in specific scenarios. Note that when the size of wireless sensor network is small, the execution time of SDR based 2BCA algorithm mainly comes from random samples generation and rescaling. In the case where parallel computation is available, this procedure can requires very little time and thus competitive to the other two algorithms.

7.7 Conclusion

This chapter considers the joint transceiver design problem in oversampling or cluster based wireless sensor network. To jointly optimize the SNR at the fusion center, the difficult original problem is decomposed into two or more subproblems and solution to

Table 7.1: MATLAB Running Time Per (Outer) Iteration(in sec.)

Dim.	Alg.	$L=5$	$L=10$	$L=20$	$L=30$	$L=40$
$K=1$	Alg.6	1.814	3.561	8.462	18.58	34.76
$N=4$	Alg.7	5.677	9.163	15.84	23.14	38.82
	Alg.8	0.067	0.380	2.603	8.344	19.38
$K=1$	Alg.6	2.175	5.413	21.97	59.23	148.5
$N=8$	Alg.7	7.488	12.21	25.58	51.19	84.17
	Alg.8	0.073	0.406	2.741	9.387	19.27
$K=3$	Alg.6	2.650	9.002	43.07	158.5	689.5
$N=4$	Alg.7	10.462	21.40	54.79	111.9	45.60
	Alg.8	1.106	2.423	6.549	14.29	26.95
$K=3$	Alg.6	7.222	32.32	536.9	—	—
$N=8$	Alg.7	19.765	50.99	173.6	59.32	85.81
	Alg.8	1.650	3.519	9.286	18.19	31.74
$K=5$	Alg.6	4.468	19.65	160.3	—	—
$K=4$	Alg.7	14.944	32.85	125.8	50.04	69.08
	Alg.8	1.455	2.989	7.749	16.63	30.41
$K=5$	Alg.6	16.442	115.8	—	—	—
$N=8$	Alg.7	32.273	121.0	80.21	134.2	201.3
	Alg.8	2.662	5.617	13.51	24.99	42.33

Note: "—" means the problem is too large to be solved.
Alg.1: SDR-2BCA alg.; Alg.2: SOCP-2BCA alg.; Alg.3: multiple BCA alg.

Table 7.2: MATLAB Running Time Per (Outer) Iteration(in sec.)

Dim.	Alg.	$K=20$	$K=30$	$K=40$
$L=2$	Alg.6	—	—	—
$N=16$	Alg.7	625.8	1.903×10^3	5.378×10^3
$M=3$	Alg.8	89.47	2.171×10^3	—

Note: "—" means the problem is too large to be solved.
Alg.1: SDR-2BCA alg.; Alg.2: SOCP-2BCA alg.; Alg.3: multiple BCA alg.

each subproblem is obtained. Convergence and complexity are carefully examined. Extensive numerical results show that the proposed algorithms provide equivalently good SNR values while have different efficiency characteristics and suitable for various system setup. As an extension of current problem, robust design and decentralized algorithms are very desirable and interesting for future study.

7.8 Appendix

7.8.1 Proof of Lemma 7.4.1

Proof. Assume that $\mathbf{X}^\star$ and $(\mathbf{Y}^\star, \nu^\star)$ are optimal solutions to (P2) and (P3) respectively, and $\mathsf{opt}(\mathrm{P2})$ and $\mathsf{opt}(\mathrm{P3})$ are optimal values of the two problems.

First we claim that $\nu^\star > 0$. This can be proved by contradiction. If $\nu^\star = 0$, then we readily obtain $\mathsf{Tr}\{\mathbf{D}_i\mathbf{Y}^\star\} = 0$, for $i = 1, \cdots, L$. This leads to $\mathsf{Tr}\{(\sum_{i=1}^L \mathbf{D}_i)\mathbf{Y}^\star\} = 0$. Since it is assumed that $\mathbf{\Sigma}_i \succ 0$, for $i \in \{1, \cdots, L\}$, it holds that $\mathbf{C}_i \succ 0$, for $i \in \{1, \cdots, L\}$. Thus $\sum_{i=1}^L \mathbf{D}_i = \mathsf{Diag}\{\mathbf{C}_1, \cdots, \mathbf{C}_L\} \succ 0$ and we obtain $\mathbf{Y}^\star = \mathbf{O}$. However this violates the constraint (7.21b), since its left hand side equals zero. Thus $\nu^\star > 0$.

If $(\mathbf{Y}^\star, \nu^\star)$ solves (P3), since $\nu^\star > 0$, it is easy to check $\mathbf{Y}^\star/\nu^\star$ is feasible for (P2) and gives an objective value of $\frac{\mathsf{Tr}\{\mathbf{A}(\mathbf{Y}^\star/\nu^\star)\}}{\mathsf{Tr}\{\mathbf{B}(\mathbf{Y}^\star/\nu^\star)+c_0\}} = \mathsf{opt}(\mathrm{P3})$. So $\mathsf{opt}(\mathrm{P3}) \leq \mathsf{opt}(\mathrm{P2})$. On the other hand, if $\mathbf{X}^\star$ solves (P2), then $\left(\frac{\mathbf{X}^\star}{\mathsf{Tr}\{\mathbf{B}\mathbf{X}^\star+c_0\}}, \frac{1}{\mathsf{Tr}\{\mathbf{B}\mathbf{X}^\star+c_0\}}\right)$ is a feasible solution to (P3) and gives objective value of $\mathsf{Tr}\{\mathbf{A}\frac{\mathbf{X}^\star}{\mathsf{Tr}\{\mathbf{B}\mathbf{X}^\star+c_0\}}\} = \mathsf{opt}(\mathrm{P2})$. So $\mathsf{opt}(\mathrm{P2}) \leq \mathsf{opt}(\mathrm{P3})$. The proof is complete. □

7.8.2 Proof of Lemma 7.4.2

Proof. First we prove that (P3) is solvable. By (7.21b) we have $0 \leq \nu \leq c_0^{-1}$, so ν is bounded. Combining (7.21c) we readily obtain $\mathsf{Tr}\{\mathbf{D}_i\mathbf{Y}\} \leq P_i\nu \leq P_i/c_0$, $i \in \{1,\cdots,L\}$, which implies $\mathsf{Tr}\{(\sum_{i=1}^{L}\mathbf{D}_i)\mathbf{Y}\} \leq (\sum_{i=1}^{L}P_i)/c_0$. Since $(\sum_{i=1}^{L}\mathbf{D}_i) = \mathsf{Diag}\{\mathbf{C}_1,\cdots,\mathbf{C}_L\} \succ 0$, this means $\mathbf{Y}$ is bounded. So the feasible region of (P3) is bounded. Obviously the feasible region of $(\mathbf{Y},\nu)$ is also closed. So (P3) has compact feasible region. Since the objective $\mathsf{Tr}\{\mathbf{AY}\}$ always takes finite values on the whole feasible region, by Weierstrass' theorem(proposition 3.2.1-(1) in [90]), (P3) is solvable.

The Lagrangian function of problem (P3) is given as

$$\begin{aligned}
&\mathcal{L}\big(\mathbf{Y},\nu,\lambda,\{\mu_i\}_{i=1}^{L}\big) \qquad (7.44)\\
=&\,\mathsf{Tr}\big\{\mathbf{YA}\big\}+\lambda\Big(1-\mathsf{Tr}\big\{\mathbf{YB}\big\}-c_0\nu\Big)-\sum_{i=1}^{L}\mu_i\Big(\mathsf{Tr}\big\{\mathbf{YD}_i\big\}-P_i\nu\Big)\\
=&\,\mathsf{Tr}\Big\{\Big[\mathbf{A}-\lambda\mathbf{B}-\sum_{i=1}^{L}\mu_i\mathbf{D}_i\Big]\mathbf{Y}\Big\}+\Big(-c_0\lambda+\sum_{i=1}^{L}P_i\mu_i\Big)\nu+\lambda.
\end{aligned}$$

By taking the supremum of Lagrangian function with respect to $\mathbf{Y} \succcurlyeq 0$ and $\nu \geq 0$, the dual function is obtained as

$$g(\lambda,\{\mu_i\}_{i=1}^{L}) = \sup_{\mathbf{Y}\succcurlyeq 0,\nu\geq 0}\mathcal{L}\big(\mathbf{Y},\nu,\lambda,\{\mu_i\}_{i=1}^{L}\big) = \lambda \qquad (7.45)$$

with the conditions $\big[\mathbf{A}-\lambda\mathbf{B}-\sum_{i=1}^{L}\mu_i\mathbf{D}_i\big] \preccurlyeq 0$ and $\big(-c_0\lambda+\sum_{i=1}^{L}P_i\mu_i\big) \leq 0$ satisfied.

So the dual problem of (P3) can be given as

$$(\mathsf{D3}): \min_{\lambda,\{\mu_i\}_{i=1}^L} \ \lambda \tag{7.46a}$$

$$\text{s.t. } \lambda\mathbf{B} + \sum_{i=1}^{L} \mu_i \mathbf{D}_i \succcurlyeq \mathbf{A}, \tag{7.46b}$$

$$c_0\lambda \geq \sum_{i=1}^{L} P_i\mu_i, \tag{7.46c}$$

$$\mu_i \geq 0, \ i = 1, \cdots, L. \tag{7.46d}$$

Next we prove that (D3) is solvable. To do this it is sufficient to show that there exists a real value γ such that the level set $\{(\lambda, \{\mu_i\}_{i=1}^L) | \lambda \leq \gamma, (\lambda, \{\mu_i\}_{i=1}^L) \in \mathsf{dom}(\mathsf{D3})\}$ is nonempty and bounded, where $\mathsf{dom}(\mathsf{D3})$ means feasible region of (D3). Here we choose $\tilde{\mu}_i = \lambda_{max}(\mathbf{A})/\lambda_{min}(\mathbf{C}_i)$ for $i = 1, \cdots, L$, where $\lambda_{max}(\cdot)$ and $\lambda_{max}(\cdot)$ represent the maximal and minimal eigenvalue of a matrix respectively. Set $\tilde{\lambda} = c_0^{-1}(\sum_{i=1}^L \tilde{\mu}_i P_i)$. By definition the constraints (7.46c) and (7.46d) are satisfied by $(\tilde{\lambda}, \{\tilde{\mu}_i\}_{i=1}^L)$. Since $\mathbf{B} \succcurlyeq 0$ and $\tilde{\lambda} \geq 0$,

$$\begin{aligned}\tilde{\lambda}\mathbf{B} + \sum_{i=1}^{L} \tilde{\mu}_i \mathbf{D}_i &\succcurlyeq \sum_{i=1}^{L} \tilde{\mu}_i \mathbf{D}_i = \mathsf{Diag}\{\tilde{\mu}_1\mathbf{C}_1, \cdots, \tilde{\mu}_L\mathbf{C}_L\} \\ &\succcurlyeq \lambda_{\max}\big(\mathbf{A}\big)\mathbf{I}_{\sum_{i=1}^L K_i N_i} \succcurlyeq \mathbf{A}. \end{aligned} \tag{7.47}$$

Thus constraint (7.46b) is also satisfied by $(\tilde{\lambda}, \{\tilde{\mu}_i\}_{i=1}^L)$. Set $\tilde{\gamma} = \tilde{\lambda}$. Combination of $\lambda \leq \tilde{\gamma}$ and the constraint (7.46c) guarantees that λ and all μ_i's are bounded. So we conclude that the level set $\{(\lambda, \{\mu_i\}_{i=1}^L) | \lambda \leq \tilde{\gamma}, (\lambda, \{\mu_i\}_{i=1}^L) \in \mathsf{dom}(\mathsf{D3})\}$ is nonempty and bounded. Invoking Weierstrass' theorem(proposition 3.2.1-(2) in [90]), (D3) is solvable. □

7.8.3 Proof of Lemma 7.4.3

Proof. Recall in definition of $\mathbf{A}_{ij}$ in (7.18b) and utilize the identity $\big(\mathbf{AB}\big) \otimes \big(\mathbf{CD}\big) = \big(\mathbf{A} \otimes \mathbf{C}\big)\big(\mathbf{B} \otimes \mathbf{D}\big)$ [89], we have

$$\begin{aligned}\mathbf{A}_{ij} &= \big(\mathbf{1}_{K_i}\mathbf{1}_{K_j}^T\big) \otimes \big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_j\big) \\ &= \big(\mathbf{1}_{K_i} \otimes \mathbf{H}_i^H\mathbf{g}\big)\big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big). \end{aligned} \tag{7.48}$$

Then the j-th block column of $\mathbf{A}$ is given as

$$\mathbf{A}_{:j} = \begin{bmatrix} \mathbf{A}_{1j} \\ \vdots \\ \mathbf{A}_{Lj} \end{bmatrix} = \begin{bmatrix} \big(\mathbf{1}_{K_1} \otimes \mathbf{H}_1^H\mathbf{g}\big)\big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big) \\ \vdots \\ \big(\mathbf{1}_{K_L} \otimes \mathbf{H}_L^H\mathbf{g}\big)\big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big) \end{bmatrix} \tag{7.49}$$

$$= \begin{bmatrix} \big(\mathbf{1}_{K_1} \otimes \mathbf{H}_1^H\mathbf{g}\big) \\ \vdots \\ \big(\mathbf{1}_{K_L} \otimes \mathbf{H}_L^H\mathbf{g}\big) \end{bmatrix} \big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big) \tag{7.50}$$

$$= \mathbf{a}\big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big). \tag{7.51}$$

The last equality utilizes the definition of $\mathbf{a}$ in (7.28). Then the matrix $\mathbf{A}$ can be represented by packing all the column blocks as follows

$$\mathbf{A} = [\mathbf{A}_{:1}, \cdots, \mathbf{A}_{:L}] \tag{7.52}$$

$$= \Big[\mathbf{a}\big(\mathbf{1}_{K_1}^T \otimes \mathbf{g}^H\mathbf{H}_1\big), \cdots, \mathbf{a}\big(\mathbf{1}_{K_L}^T \otimes \mathbf{g}^H\mathbf{H}_L\big)\Big] \tag{7.53}$$

$$= \mathbf{a}\Big[\big(\mathbf{1}_{K_1}^T \otimes \mathbf{g}^H\mathbf{H}_1\big), \cdots, \big(\mathbf{1}_{K_L}^T \otimes \mathbf{g}^H\mathbf{H}_L\big)\Big] \tag{7.54}$$

$$= \mathbf{a}\mathbf{a}^H. \tag{7.55}$$

The proof is complete. □

7.8.4 Proof of Lemma 7.4.4

Proof. By the i-th power constraint (7.19b) we have

$$\lambda_{min}\big(\mathbf{C}_i\big)\|\mathbf{f}_i\|_2^2 \leq \mathbf{f}_i^H\mathbf{C}_i\mathbf{f}_i \leq P_i, \tag{7.56}$$

which implies

$$\|\mathbf{f}_i\|_2 \leq \sqrt{\frac{P_i}{\lambda_{min}(\mathbf{C}_i)}},\ i = 1, \cdots, L. \tag{7.57}$$

By Cauchy-Schwarz inequality the numerator $\mathbf{f}^H\mathbf{A}\mathbf{f}$ of SNR is bounded as

$$\mathbf{f}^H\mathbf{A}\mathbf{f} = \big|\mathbf{a}^H\mathbf{f}\big|^2 \leq \Big|\sum_{i=1}^{L}\big|\mathbf{f}_i^H\big(1_{K_i}\otimes\mathbf{H}_i^H\mathbf{g}\big)\big|\Big|^2 \tag{7.58}$$

$$\leq \Big|\sum_{i=1}^{L}\big\|\mathbf{f}_i\big\|_2\big\|\mathbf{1}_{K_i}\otimes\mathbf{H}_i^H\mathbf{g}\big\|_2\Big|^2 \tag{7.59}$$

$$= \left(\sum_{i=1}^{L}\sqrt{\frac{P_i}{\lambda_{min}\big(\mathbf{C}_i\big)}}K_i\big\|\mathbf{H}_i^H\mathbf{g}\big\|_2\right)^2, \tag{7.60}$$

where the above first inequality utilizes 7.4.3. Combining the fact that $\mathbf{f}^H\mathbf{B}\mathbf{f} + c_0 \geq c_0$, the upper bound in lemma is proved. □

7.8.5 Proof of Theorem 7.4.2

Proof. Since each update of $\{\mathbf{F}_i\}_{i=1}^L$ or $\mathbf{g}$ is obtained by solving a maximization problem, SNR monotonically increases. SNR is also bounded. In fact since the SNR is invariant to scaling of $\mathbf{g}$ we can assume that $\|\mathbf{g}\|_2 = 1$. According to (7.57) in the proof of lemma 7.4.4, $\mathbf{F}_i$ is bounded for all $i = 1, \cdots, L$. Thus the numerator of SNR is bounded above and the denominator of SNR is bounded away from zero, so SNR should be bounded. Consequently the objective value sequence by algorithms 6 or 7 converges since it is monotonically increasing and bounded.

Since $\{\mathbf{F}_i\}_{i=1}^L$ are bounded, by Bolzano-Weierstrass theorem [91] there exists a sequence $\{j_k\}_{k=1}^\infty$ such that $\left\{\{\mathbf{F}_i^{(j_k)}\}_{i=1}^L\right\}_{k=1}^\infty$ converges. Since $\mathbf{g}^{(j_k)}$ is updated by (7.8) which is a continuous function of $\{\mathbf{F}_i^{(j_k)}\}_{i=1}^L$, thus the sequence $\left\{\left(\{\mathbf{F}_i^{(j_k)}\}_{i=1}^L, \mathbf{g}^{(j_k)}\right)\right\}_{k=1}^\infty$ also converges. The existence of limit points of the solution sequence is proved.

The feasible region of (P0) is a Cartesian product $\mathcal{X}_1 \times \mathcal{X}_2$ with $\mathcal{X}_1 \triangleq \left\{\{\mathbf{F}_i\}_{i=1}^L \middle| (7.7b) \right.$ is satisfied for$\left. i = 1, \cdots, L\right\}$ and $\mathcal{X}_2 \triangleq \mathbb{C}^{M\times 1}\backslash\{\mathbf{0}\}$. Corollary 2 in [64] states that any limit point of solution sequence generated by 2-block coordinate ascent method is stationary. It should be noted that this conclusion is obtained under the assumption that the objective function is continuously differentialbe on feasible region and each block feasible region(each term in the Cartesian product) is nonempty, closed and convex set. Unfortunately the problem (P0) does not satisfy this assumption since $\mathcal{X}_2$ is nonconvex and not closed. In the following we will show that conclusion in [64] still applies to our problem after appropriate adjustment in its proof.

First we assert that the solution sequence always has nonzero $\mathbf{g}$, i.e. $\mathbf{g}^{(k)} \neq 0$ for all $k = 0, 1, \cdots$. Since algorithms 6 or 7 starts from $\left(\{\mathbf{F}_i^{(0)}\}_{i=1}^L, \mathbf{g}^{(0)}\right)$ with $\mathbf{g}^{(0)} \neq \mathbf{0}$, the as-

sertion holds for $k = 0$ and $\mathsf{SNR}\big(\{\mathbf{F}_i^{(0)}\}_{i=1}^L, \mathbf{g}^{(0)}\big) > 0$. Assume that $m \geq 1$ is the smallest integer such that $\mathbf{g}^{(m)} = \mathbf{0}$, then according to (7.8) $\big(\sum_{i=1}^L \mathbf{H}_i \mathbf{F}_i^{(m)} \mathbf{1}_{K_i}\big) = \mathbf{0}$. Notice $\mathbf{g}^{(m-1)} \neq \mathbf{0}$. This readily implies $\mathsf{SNR}\big(\{\mathbf{F}_i^{(m)}\}_{i=1}^L, \mathbf{g}^{(m-1)}\big) = 0 < \mathsf{SNR}\big(\{\mathbf{F}_i^{(0)}\}_{i=1}^L, \mathbf{g}^{(0)}\big)$, which contradicts the increasing monotonicity of SNR.

Then we assert that any limit point of solution sequence has nonzero $\bar{\mathbf{g}}$. By contradiction we assume that the subsequence $\Big\{\big(\{\mathbf{F}_i^{(j_k)}\}_{i=1}^L, \mathbf{g}^{(j_k)}\big)\Big\}_{k=1}^\infty$ converges to $(\{\bar{\mathbf{F}}_i\}_{i=1}^L, \bar{\mathbf{g}})$ with $\bar{\mathbf{g}} = \mathbf{0}$. Then by (7.8) $\Big\{\sum_i^L \mathbf{H}_i \mathbf{F}_i^{(j_k)} \mathbf{1}_{K_i}\Big\}_{k=1}^\infty \to \mathbf{0}$. By rescaling each $\mathbf{g}^{(j_k)}$ to $\hat{\mathbf{g}}^{(j_k)}$ such that $\|\hat{\mathbf{g}}^{(j_k)}\|_2 = 1$ for all $k = 1, 2, \cdots$, we actually construct another solution sequence which is also generated by 2-block coordinate ascent method, since scaling of $\mathbf{g}$ does not change the SNR value. Now for this new solution sequence $\Big\{\big(\{\mathbf{F}_i^{(j_k)}\}_{i=1}^L, \hat{\mathbf{g}}^{(j_k)}\big)\Big\}_{k=1}^\infty$, since $\Big\{\sum_i^L \mathbf{H}_i \mathbf{F}_i^{(j_k)} \mathbf{1}_{K_i}\Big\}_{k=1}^\infty \to \mathbf{0}$ while $\{\hat{\mathbf{g}}^{(j_k)}\}_{k=1}^\infty$(consequently the denominator of SNR) is bounded away from zero, we have $\mathsf{SNR}\big(\{\mathbf{F}_i^{(j_k)}\}_{i=1}^L, \hat{\mathbf{g}}^{(j_k)}\big) \to 0$, which again contradicts the increasing monotonicity of SNR sequence.

In [64] the closedness assumption of $\mathcal{X}_2$ is implicitly invoked in its proposition 2 to ensure that any limit point of solution sequence is feasible. Through the above proof we can see that this result holds true thus proposition 2 in [64] applies to our problem.

The convexity assumption of $\mathcal{X}_2$ is explicitly utilized in [64] in its proof of proposition 3. Here we identify the notations i, $\mathbf{x}_{i+1}$, $\mathcal{X}_{i+1}$ and $\mathbf{w}(k, i)$ used in the original proof of proposition 3 in [64] as 1, $\mathbf{g}$, $\mathcal{X}_2$ and $\big(\{\mathbf{F}_i^{(k+1)}\}_{i=1}^L, \mathbf{g}^{(k)}\big)$ respectively in our case.

According to the proof in [64], the convexity of $\mathcal{X}_2$ guarantees that $\mathbf{g}^{(k)} + \alpha_2^{(k)} \mathbf{d}_2^{(k)} \in \mathcal{X}_2$ such that Armijo-type line search algorithm's properties can be applied, where $\mathbf{d}_2^{(k)} =$

$\tilde{\mathbf{g}} - \mathbf{g}^{(k)}$ with $\tilde{\mathbf{g}} \in \mathcal{X}_2$ and

$$\partial^H_{\mathbf{g}^{*(k)}} \mathsf{SNR}\big(\{\mathbf{F}_i^{(k+1)}\}_{i=1}^L, \mathbf{g}^{(k)}\big)\big(\tilde{\mathbf{g}} - \mathbf{g}^{(k)}\big) < 0 \tag{7.61}$$

((7.61) corresponds to the inequality $\nabla_{i+1} f(\mathbf{w}(k,i))^T d_{i+1}^k < 0$ in the original proof of proposition 3 in [64], which lies under equation (11) and is not labeled).

Now we show that the fact $\mathbf{g}^{(k)} + \alpha_2^{(k)}\mathbf{d}_2^{(k)} \in \mathcal{X}_2$ still holds for our problem although our $\mathcal{X}_2$ is nonconvex. By contradiction assume that $\mathbf{g}^{(k)} + \alpha_2^{(k)}\mathbf{d}_2^{(k)} \notin \mathcal{X}_2$, i.e.

$$\mathbf{g}^{(k)} + \alpha_2^{(k)}\mathbf{d}_2^{(k)} = \mathbf{0}. \tag{7.62}$$

This is actually impossible. By substituting $\mathbf{d}_2^{(k)} = \tilde{\mathbf{g}} - \mathbf{g}^{(k)}$ into (7.62) we have

$$(\alpha_2^{(k)} - 1)\mathbf{g}^{(k)} = \alpha_2^{(k)}\tilde{\mathbf{g}}. \tag{7.63}$$

As a result of Armijo-type line search algorithm(refer to (3) and proposition 1 in [64]), $(\alpha_2^{(k)} - 1) \in (0, 1]$. If $\alpha_2^{(k)} = 1$, then $\tilde{\mathbf{g}} = \mathbf{0}$, which contradicts the fact $\tilde{\mathbf{g}} \in \mathcal{X}_2$. If $(\alpha_2^{(k)} - 1) < 1$, then $\tilde{\mathbf{g}} = \frac{(\alpha_2^{(k)}-1)}{\alpha_2^{(k)}}\mathbf{g}^{(k)}$. This is also impossible since SNR is invariant to scaling of $\mathbf{g}$ and thus $\partial^H_{\mathbf{g}^{(k)*}} \mathsf{SNR}\big(\{\mathbf{F}_i^{(k+1)}\}_{i=1}^L, \mathbf{g}^{(k)}\big)\big(\tilde{\mathbf{g}} - \mathbf{g}^{(k)}\big) = 0$, which contradicts the fact (7.61). Thus the proposition 3 in [64] also stands for our problem.

As a direct implication of proposition 2 and 3, the corollary 2 in [64] holds true and thus any limit point provided by algorithm 6 or 7 is stationary point of (P0). □

7.8.6 Proof of Theorem 7.5.1

Proof. Since the problem $(\mathsf{P}8^i_\alpha)$ is a quadratic problem with one quadratic constraint and is obviously strictly feasible, the result of Appendix B.1 in [68] is valid to invoke, which states that $(\mathsf{P}8^i_\alpha)$ has the following relaxation

$$(\mathsf{P}9^i_\alpha) \min_{\mathbf{X},\mathbf{x}} . \ \mathsf{Tr}\big\{\big[\alpha\mathbf{B}_i - \mathbf{A}_{ii}\big]\mathbf{X}\big\} - 2\mathsf{Re}\big\{\mathbf{q}_i^H\mathbf{x}\big\} + (\alpha d_i - c_i), \tag{7.64a}$$

$$\text{s.t. } \mathsf{Tr}\big\{\mathbf{C}_i\mathbf{X}\big\} - P_i \leq 0, \tag{7.64b}$$

$$\begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^H & 1 \end{bmatrix} \succeq 0. \tag{7.64c}$$

with $\mathsf{opt}(\mathsf{P}9^i_\alpha) = \mathsf{opt}(\mathsf{P}8^i_\alpha)$. We replace the variables $(\mathbf{X}, \mathbf{x})$ in $(\mathsf{P}9^i_\alpha)$ by one matrix variable $\widetilde{\mathbf{X}}$ and rewrite it into a SDP form

$$(\mathsf{P}10^i_\alpha) \min_{\widetilde{\mathbf{X}}} . \ \mathsf{Tr}\big\{\mathbf{P}_1\widetilde{\mathbf{X}}\big\}, \tag{7.65a}$$

$$\mathsf{s.t.}\ \mathsf{Tr}\big\{\mathbf{P}_2\widetilde{\mathbf{X}}\big\} \leq P_i, \tag{7.65b}$$

$$\mathsf{Tr}\big\{\mathbf{P}_3\widetilde{\mathbf{X}}\big\} = 1, \tag{7.65c}$$

with the parameter matrices being defined as

$$\mathbf{P}_1 \triangleq \begin{bmatrix} \alpha\mathbf{B}_i - \mathbf{A}_{ii} & -\mathbf{q}_i \\ -\mathbf{q}_i^H & \alpha d_i - c_i \end{bmatrix}, \mathbf{P}_2 \triangleq \begin{bmatrix} \mathbf{C}_i & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \mathbf{P}_3 \triangleq \begin{bmatrix} \mathbf{O} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}.$$

Since $\mathbf{C}_i \succ 0$, the feasible set of $(\mathsf{P}10^i_\alpha)$ is bounded. Obviously the objective of $(\mathsf{P}10^i_\alpha)$ takes finite value over the feasible set, so $(\mathsf{P}10^i_\alpha)$ is solvable by WeierstrassâĂŹs

theorem(proposition 3.2.1-(1) in [90]).

Assume that $\widetilde{\mathbf{X}}^\star$ is one optimal solution. Obviously $\widetilde{\mathbf{X}}^\star$ is non-zero(otherwise constraint $\mathbf{P}_3\widetilde{\mathbf{X}}^\star = 1$ would fail). Since $\widetilde{\mathbf{X}}^\star$ has dimension $K_iN_i + 1 \geq 3$, evoking theorem 2.2 of [86], we can obtain a vector $\widetilde{\mathbf{x}}$ such that $\mathsf{Tr}\{\mathbf{P}_j\widetilde{\mathbf{X}}^\star\} = \mathsf{Tr}\{\mathbf{P}_j\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^H\}$ for $j = 1, 2, 3$. Denote $\widetilde{\mathbf{x}} = [\widetilde{\mathbf{x}}_1^T, \tilde{x}_2]^T$. Notice that $\tilde{x}_2$ is nonzero(otherwise the constraint $\mathsf{Tr}\{\mathbf{P}_3\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^H\} = |\tilde{x}_2|^2 = 1$ would fail). Define $\widehat{\mathbf{x}} \triangleq [\widetilde{\mathbf{x}}_1^T/\tilde{x}_2, 1]^T$, it is easy to check that

$$\begin{aligned} f_0\big(\frac{\widetilde{\mathbf{x}}_1}{\tilde{x}_2}\big)&=\mathsf{Tr}\{\mathbf{P}_1\widehat{\mathbf{x}}\widehat{\mathbf{x}}^H\}=\mathsf{Tr}\{\mathbf{P}_1\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^H\}=\mathsf{Tr}\{\mathbf{P}_1\widetilde{\mathbf{X}}^\star\}=\mathsf{opt}(\mathsf{P}10^i_\alpha);\\ f_1\big(\frac{\widetilde{\mathbf{x}}_1}{\tilde{x}_2}\big)&=\mathsf{Tr}\{\mathbf{P}_2\widehat{\mathbf{x}}\widehat{\mathbf{x}}^H\}=\mathsf{Tr}\{\mathbf{P}_2\widetilde{\mathbf{x}}\widetilde{\mathbf{x}}^H\}=\mathsf{Tr}\{\mathbf{P}_2\widetilde{\mathbf{X}}^\star\}\leq P_i. \end{aligned} \tag{7.67}$$

where $f_0(\cdot)$ denotes the objective function of $(\mathsf{P}8^i_\alpha)$ and $f_1(\mathbf{x}) \triangleq \mathbf{x}^H\mathbf{C}_i\mathbf{x}$. The above two equations imply that $\widetilde{\mathbf{x}}_1/\tilde{x}_2$ is an optimal solution to $(\mathsf{P}8^i_\alpha)$ since which gives optimal value $\mathsf{opt}(\mathsf{P}10^i_\alpha)$ and is feasible. □

7.8.7 Proof of Theorem 7.5.2

Proof. When $K_i = 1$, the covariance matrices $\mathbf{\Sigma_s}$ and $\{\mathbf{\Sigma}_i\}_{i=1}^L$ become scalars σ_s^2 and $\{\sigma_i^2\}_{i=1}^L$ respectively and we have

$$\mathbf{A}_{ii} = \mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i, \quad \mathbf{B}_i = \sigma_i^2\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i, \tag{7.68a}$$

$$\mathbf{C}_i = (\sigma_i^2{+}\sigma_s^2)\mathbf{I}_{N_i}, \tag{7.68b}$$

$$\mathbf{A}_{ij} = \mathbf{1}_{K_j}^T \otimes \big(\mathbf{H}_i^H\mathbf{g}\mathbf{g}^H\mathbf{H}_i\big) = \mathbf{H}_i^H\mathbf{g}\big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big), \tag{7.68c}$$

$$\mathbf{q}_i = \sum_{j\neq i}\mathbf{A}_{ij}\mathbf{f}_j = \mathbf{H}_i^H\mathbf{g}\Big[\sum_{j\neq i}\big(\mathbf{1}_{K_j}^T \otimes \mathbf{g}^H\mathbf{H}_j\big)\mathbf{f}_j\Big]. \tag{7.68d}$$

To simplify the following discussion, we introduce the notations

$$\beta_i \triangleq \sum_{j\neq i} \mathbf{f}_j^H \Big(\mathbf{1}_{K_j} \otimes \mathbf{H}_j^H \mathbf{g}\Big),\ \bar{P}_i \triangleq \frac{P_i}{\sigma_i^2 + \sigma_s^2}. \tag{7.69}$$

Then the problem $(\mathsf{P8}_\alpha^i)$ in (7.40) is expressed as

$$\begin{aligned}(\mathsf{P11}_\alpha^i): \min_{\mathbf{f}_i}.\ & \big(\alpha\sigma_i^2 - 1\big)\mathbf{f}_i^H \mathbf{H}_i^H \mathbf{g}\mathbf{g}^H \mathbf{H}_i \mathbf{f}_i - 2\mathsf{Re}\big\{\beta_i \mathbf{g}^H \mathbf{H}_i \mathbf{f}_i\big\} \\ & + \big(\alpha d_i - c_i\big), \qquad (7.70a)\\ \text{s.t. } & \|\mathbf{f}_i\|^2 \leq \bar{P}_i. \qquad (7.70b)\end{aligned}$$

The key observation is that the quadratic matrix $\mathbf{H}_i^H \mathbf{g}\mathbf{g}^H \mathbf{H}_i$ in the objective function has rank one and thus an eigenvalue decomposition as follows

$$\mathbf{H}_i^H \mathbf{g}\mathbf{g}^H \mathbf{H}_i = \mathbf{U} \begin{bmatrix} \mathbf{g}^H \mathbf{H}_i \mathbf{H}_i^H \mathbf{g} & \\ & \mathbf{O}_{(N_i-1)\times(N_i-1)} \end{bmatrix} \mathbf{U}^H, \tag{7.71}$$

with $\mathbf{U} \triangleq \big[\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_i}\big]$ being eigenvectors of $\mathbf{H}_i^H \mathbf{g}\mathbf{g}^H \mathbf{H}_i$. The first eigenvector $\mathbf{u}_1$ corresponds to the unique nonzero eigenvalue and the other eigenvectors span the null space of $\mathbf{H}_i^H \mathbf{g}\mathbf{g}^H \mathbf{H}_i$. In other words, we have

$$\mathbf{u}_1 = \frac{\mathbf{H}_i^H \mathbf{g}}{\|\mathbf{H}_i^H \mathbf{g}\|}, \quad \mathbf{u}_j^H \mathbf{H}_i^H \mathbf{g} = 0, \quad j = \{2, \cdots, N_i\} \tag{7.72}$$

Since $\{\mathbf{u}_i\}_{i=1}^L$ is an orthonormal basis, $\mathbf{f}_i$ can be represented as $\mathbf{f}_i = \mathbf{U}\tau = \sum_{j=1}^{N_i} \mathbf{u}_j \tau_j$ with vector τ being the coordinates in terms of basis $\{\mathbf{u}_j\}_{j=1}^{N_i}$.

By (7.72) the objective of $(\mathsf{P11}_\alpha^i)$ is independent of $\{\tau_j\}_{j=2}^{N_i}$. To save power, we

should set all $\{\tau_j\}_{j=2}^{N_i}$ as zero, which means $\mathbf{f}_i = \tau_1 \frac{\mathbf{H}_i^H \mathbf{g}}{\|\mathbf{H}_i^H \mathbf{g}\|_2}$. Thus, the problem $(\mathsf{P11}_\alpha^i)$ boils down to the following problem with respect to one complex scalar τ_1

$$(\mathsf{P12}_\alpha^i) : \min_{\tau_1} . \; g(\tau_1) \triangleq (\alpha\sigma_i^2 - 1)\|\mathbf{H}_i^H \mathbf{g}\|_2^2 |\tau_1|^2 \tag{7.73a}$$

$$-2\|\mathbf{H}_i^H \mathbf{g}\|_2 \mathsf{Re}\{\beta_i \tau_1\} + \big(\alpha d_i - c_i\big),$$

$$\text{s.t. } |\tau_1|^2 \leq \bar{P}_i. \tag{7.73b}$$

Based on the sign of $(\sigma_i^2 - \alpha)$, the problem $(\mathsf{P12}_\alpha^i)$ can be tackled in the following three cases:

CASE (I): $\alpha = \sigma_i^{-2}$. In this case, the objective function in $(\mathsf{P12}_\alpha^i)$ degenerates to an affine function

$$g(\tau_1) = -2\sigma_i^2 \|\mathbf{H}_i^H \mathbf{g}\|_2 \mathsf{Re}\{\beta_i \tau_1\} + (\alpha d_i - c_i). \tag{7.74}$$

By the Cauchy-Schwarz inequality, the optimal $\tau_1^\star$ and minimum objective is obtained as

$$\tau_1^\star = \frac{\sqrt{\bar{P}_i}\beta_i^*}{|\beta_i|}; g(\tau_1^\star) = -2\sigma_i^2 \|\mathbf{H}_i \mathbf{g}\|_2 |\beta_i| \sqrt{\bar{P}_i} + \big(\alpha d_i - c_i\big); \tag{7.75}$$

When $\alpha \neq \sigma_i^{-2}$ by denoting

$$\zeta_i \triangleq \big(\alpha d_i - c_i\big) - \frac{|\beta_i|^2}{\alpha\sigma_i^2 - 1}, \tag{7.76}$$

the objective function is equivalently written as

$$g(\tau_1)=(\alpha\sigma_i^2-1)\|\mathbf{H}_i^H\mathbf{g}\|_2^2\left|\tau_1-\frac{\beta_i^*}{\|\mathbf{H}_i^H\mathbf{g}\|_2(\alpha\sigma_i^2-1)}\right|^2+\zeta_i, \tag{7.77}$$

CASE (II): $\alpha > \sigma_i^{-2}$. To minimize $g(\tau_1)$, τ_1 should be along the direction of β_i^*. Depending on whether the zero point of the absolute term in (7.77) satisfies the power constraint, two subcases are examined:

i) If $\frac{|\beta_i^*|}{\|\mathbf{H}_i^H\mathbf{g}\|_2(\alpha\sigma_i^2-1)} \leq \sqrt{\bar{P}_i}$, the optimum is given as

$$\tau_1^\star=\frac{\beta_i^*}{\|\mathbf{H}_i^H\mathbf{g}\|_2(\alpha\sigma_i^2-1)};\quad g(\tau_1^\star)=\zeta_i; \tag{7.78}$$

ii) If $\frac{|\beta_i^*|}{\|\mathbf{H}_i^H\mathbf{g}\|_2(\alpha\sigma_i^2-1)} > \sqrt{\bar{P}_i}$, the optimum is given as

$$\begin{aligned}\tau_1^\star &= \frac{\sqrt{\bar{P}_i}\beta_i^*}{|\beta_i|}; \\ g(\tau_1^\star)&=(\alpha\sigma_i^2-1)\|\mathbf{H}_i^H\mathbf{g}\|_2^2\left|\sqrt{\bar{P}_i}-\frac{|\beta_i|}{\|\mathbf{H}_i^H\mathbf{g}\|_2(\alpha\sigma_i^2-1)}\right|^2+\zeta_i\end{aligned} \tag{7.79}$$

CASE (III): $\alpha < \sigma_i^{-2}$. Still τ_1 should be along the direction of β_i^*, and takes full power. At this time the optimum is literally identical with (7.79) in the above.

Remember that optimal $\mathbf{f}_i^\star$ to $(\mathrm{P}8_\alpha^i)$ is obtained by $\mathbf{f}_i^\star = \tau_1^\star\frac{\mathbf{H}_i^H\mathbf{g}}{\|\mathbf{H}_i^H\mathbf{g}\|_2}$, the proof is complete. □

Chapter 8

Conclusion

This dissertation research focuses on efficient signal encoding and decoding schemes in both point-to-point and network communication systems. It is briefly summarized in the follows.

In chapter 2 we discuss the performance limits of linear analog codes under AWGN channel using ML and LMMSE decoding methods. The performance bound has been obtained and the optimal linear analog codes have been obtained. It is shown that these two different decoding methods lead to identical linear encoding schemes— unitary codes. Extensive numerical results verify what we have found.

Chapter 3 researches a kind of nonlinear analog encoding scheme based on chaotic dynamic systems. Under the general AWGN channel setup, various optimal and sub-optimal decoding algorithms, including MMSE, ML and ML-MMSE algorithms, have been developed. Based on the Cramer-Rao bound(CRB) for the baker's encoding systems and analysis of its performance, two improving schemes are proposed—mirrored

baker's dynamic systems and one input baker's system. These two improvements effectively depress the threshold effect of the original system and significantly outperform the classical tent's map chaotic encoding system.

Chapter 4 considers the precoding design for single sensor with single antenna by exploiting signal space diversity. By analyzing pairwise error probability(PEP), we discuss precoder design criteria. Additionally PNC algorithm is proposed which presents a good compromise between the decoding complexity and bit error rate(BER) performance.

Chapter 5 focuses on transceiver design problem minimizing MSE in a centralized wireless sensor network. Since the original problem is difficult, we adopt BCD method to attack it. A 2-BCD algorithm is first proposed, which has an easy subproblem of MMSE receiving and one much harder convex subproblem, which can be turned into a SOCP problem. Based on that another multiple block BCD algorithm has been proposed by further decomposing the harder subproblem into more blocks with closed solution to each block obtained. The convergence and complexity of the proposed algorithms have been discussed.

In chapter 6, we consider the problem of maximizing MI in the centralized wireless sensor network. Inspired by the WMMSE method, we introduce two complicating intermediate variables—and perform the BCD method to decompose the MI problem into three subproblems. Two subproblems have closed form solutions and the other one is similar to one in 2-BCD decomposition in MSE problem. The Karush-Kuhn-Tucker(KKT) conditions have been examined and we manage to prove that the solutions of this 3-BCD algorithm are KKT points. Convergence has also been carefully examined.

In chapter 7, we consider the problem of maximizing signal to noise ratio(SNR) in centralized wireless sensor network. Still utilizing the block coordinate descent method, we decompose the original problem into subproblems optimizing individual transmitter or FC receiver alone. For the receiver optimization, the subproblem can be easily solved by generalized eigenvalue decomposition. For the transmitter optimization subproblem, by use of S-lemma and rank-one decomposition, it can be solved semidefinite programming rank-one decomposition. Numerical results are provided to verify our proposed algorithms.

Bibliography

[1] T. G. Marshall, Jr. “Real number transform and convolutional codes,” *Proc. 24th Midwest Symp. Circuits Sys.*, Albuquerque, NM, USA, June, 1981.

[2] J. K. Wolf, “Analog codes,” *IEEE Intl. Conf. Comm,* Boston, MA, USA, June, 1983.

[3] J. K. Wolf, “Redundancy, the discrete Fourier transform, and impulse noise cancellation,” *IEEE Trans. Comm.*, vol. 31, no. 3, pp. 458-461, March 1983.

[4] J. L. Wu and J. Shiu, “Discrete cosine transform in error control coding”, *IEEE Trans. Comm.*, vol. 43, no. 5, pp. 1857-1861, May 1995.

[5] O. Tirkkonen, and A. Hottinen, “Square-matrix embeddable space-time block codes for complex signal constellations,” *IEEE Trans. Info. Theory*, vol. 48, no. 2, pp. 384-395, Feb. 2002.

[6] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, R. Urbanke, “Systematic design of unitary space-time constellations,” vol. 46, no. 6, *IEEE Trans. Info. Theory*, pp. 1962-1973, Sep. 2000.

[7] G. Rath and C. Guillemot, "Characterization of a class of error correcting frames for robust signal transmission over wireless communication channels", *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 2, pp. 229-241, 2005.

[8] W. Henkel, "Analog codes for peak-to-average ratio reduction," Proc. of *3rd ITG Conf. Source and Channel Coding*, Munich, Germany, Apr. 2000.

[9] Z. Wang and G. Giannakis, "Complex-field coding for OFDM over fading wireless channels," *IEEE Trans Info Theory*, vol. 49, no. 3, pp. 707-720, March 2003.

[10] K. H. Lee and D. P. Petersen, "Optimal linear coding for vector channels", *IEEE Trans. Commu.*, vol. 24, no. 12, pp. 1283-1290, Dec. 1976.

[11] S. Ohno, and G. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels", *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2138-2145 Sept. 2004.

[12] C. E. Shannon, "Communication in the presence of noise," in *Proceedings of IRE*, vol. 37, no. 1, pp. 10-21, 1949.

[13] V.A.Kotel'nikov, *The Theory of Optimum Noise Immunity*, McGraw-Hill, 1959.

[14] J. M. Wozencraft, I. M. Jacobs, *Principles of Communication Engineering*, John Wiley & Sons, 1965.

[15] T. J. Goblick, "Theoretical limitation on the transmission of data from analog sources," *IEEE Trans. Inform. Theory,* vol. 11, no. 4, pp. 558-566, Oct. 1965.

[16] J. Ziv, "The behavior of analog communication systems," *IEEE Trans. Inform. Theory,* vol. 16, no. 5, pp. 587-594, Sep. 1970.

[17] K. H. Lee, D P Petersen, “Optimal linear coding for vector channels,” *IEEE Trans. Commun.,* vol. 24, no. 12, pp. 1283-1290, Dec. 1976.

[18] Y. Liu, J. Li, K. Xie, “Analysis of linear channel codes with continuous code space,” *46th Annual Conference on Information Sciences and Systems*, Princeton, USA, March, 2012

[19] A. Fuldseth, Robust subband video compression for noisy channels with multilevel signaling, Dissertation, Norwegian University of Science and Technology, 1997.

[20] S.-Y. Chung, On the construction of some capacity-approaching coding schemes, Dissertation, Massachusetts Institute of Technology, 2000.

[21] V. A. Vaishampayan, S. I. R. Costa, “Curves on a sphere shift-map dynamics and error control for continuous alphabet sources,” *IEEE Trans. Inform. Theory,* vol. 49, no. 7, pp. 1658-1672, July 2003.

[22] X. Cai, J. W. Modestino, “Bandwidth expansion Shannon mapping for analog error-control coding,” in *40th Annual Conference on Information Sciences and Systems*, Princeton, USA, March, 2006

[23] Y. Hu, J. Garcia-Frias, M. Lamarca, “Analog joint source-channel coding using non-linear curves and MMSE decoding,” *IEEE Trans. Commun.,* vol.59, no. 11, pp. 3016-3026, Sep. 2011.

[24] G. Brante, R. D. Souza, J. Garcia-Frias, “Spatial diversity using analog joint source channel coding in wireless channels,” *IEEE Trans. Commun.,* vol. 61, no. 1, pp. 301-311, Feb. 2013.

[25] F. Hekland, P. Floor, T. Ramstad, "Shannon-KotelâĂŹnikov mappings in joint source-channel coding," *IEEE Trans. Commun.,* vol. 57, no. 11, pp. 94-105, Jan. 2009.

[26] P. A. Floor, T. A. Ramstad, "Shannon-KotelâĂŹnikov mappings for analog point-to-point communications," http://arxiv.org/abs/0904.1538

[27] H. C. Papadopoulus, G. W. Wornell, "Maximum likelihood estimation of a class of chaotic signals," *IEEE Trans. Inform. Theory,* vol. 41, no.1, pp. 312-317, Jan. 1995.

[28] B. Chen, G. W. Wornell, "Analog error-correction codes based on chaotic dynamical systems," *IEEE Trans. Commun.,* vol. 46, no. 7, pp. 881-890, July 1998.

[29] S. M. Kay, "Asymptotic maximum likelihood estimator performance for chaotic signals in noise,", *IEEE Trans. Signal Process.,* vol. 43, no. 4, pp. 1009-1012, Apr. 1995.

[30] I. Hen, N. Merhav, "On the threshold effect in the estimation of chaotic sequences," *IEEE Trans. Inform. Theory,* vol. 50, no. 11, pp. 2894-2904, Nov. 2004.

[31] I. Rosenhouse, A. J. Weiss, "Combined analog and digital error-correcting codes for analog information sources," *IEEE Trans. Commun.,* vol. 55, no. 11, pp. 2073-2083, Nov. 2007.

[32] H. V. Poor, *An Introduction to Signal Detection and Estimation, 2nd ed.,* Springer, 1998.

[33] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I, Estimation Theory*, Prentice Hall, 1993.

[34] S. Haykin, *Adaptive Filter Theory, 4th ed.,* Prentice Hall, 2002.

[35] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* John Wiely & Sons, 1991.

[36] I.F. Akyildiz, T. Melodia and K. R. Chowdury, "A survey on wireless multimedia sensor networks", *Computer Networks*, vol. 51, pp 921-960, Dec. 2007.

[37] Y. Liu, J. Li, K. Xie, J. Hou, "Wireless data sensing and transmission through analog codes," *Conference on Information Sciences and Systems (CISS)*, Princeton, N.J., U.S.A., Mar. 2012.

[38] Y. Liu, J. Li and K. Xie, "Analysis of linear channel codes with continuous code space," *Conference on Information Sciences and Systems (CISS)*, Princeton, N.J., U.S.A., Mar. 2012.

[39] Y. Liu, X. Lu, J. Li, "Reliable Signal Transmission in Wireless Sensor Networks with Zero Bandwidth Expansion," *IEEE International Conference on Communications (ICC)*, Budapest , Jun. 2013.

[40] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,"*Math. Comput.*, vol. 44, pp. 463-471, Apr. 1985.

[41] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. expected complexity," *IEEE Trans. Signal Proc.*, vol. 53, no. 8, pp. 2806-2818, Aug. 2005.

[42] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice codes decoder for space-time codes," *IEEE Commun. Lett.*, vol. 4, no. 5, pp. 161-163, May 2000.

[43] J. W. Craig, “A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations”, *IEEE MILCOM’91, Conf. Rec.*, Mclean, VA, USA, 1991.

[44] M. K. Simon and M.S. Alouini, “Digital communication over fading channels: a unified approach to performance analysis”, New York, NY: John Wiley and Sons, 2000.

[45] J. Boutros and E. Viterbo, “Signal space diversity: a power and bandwidth efficient diversity technique for the rayleigh fading channel,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 1453-1467, July 1998.

[46] V. Tarokh, N. Seshadri, and R. Calderbank, “Space-time codes for high data rate wireless communication: performance criterion and code construction,” *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.

[47] Y. Xin, Z. Wang, and G. B. Giannakis, “Space-time diversity systems based on linear constellation precoding,” *IEEE Trans. Wireless. Commun.*, vol. 2, no. 2, pp. 294-309, Mar. 2003.

[48] Y. Zhu, E. Song, J. Zhou, and Z. You, “Optimal dimensionality reduction of sensor data in multisensor estimation fusion,” *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1631-1639, May 2005.

[49] J. Fang and H. Li, “Optimal/near-optimal dimensionality reduction for distributed estimation in homogeneous and certain inhomogeneous scenarios,” *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4339-4353, 2010.

[50] J. Fang and H. Li, "Power constrained distributed estimation with cluster-based sensor collaboration," *IEEE Trans. Wireless Commun.,* vol. 8, no. 7, pp. 3822-3832, July 2009.

[51] I. D. Schizas, G. B. Giannakis, and Z.-Q. Luo, "Distributed estimation using reduced-dimensionality sensor observations," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4284-4299, Aug. 2007.

[52] J. Xiao, S. Cui, Z. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Process.,* vol. 56, no. 2, pp. 757-770, Feb. 2008.

[53] A. S. Behbahani, A. M. Eltawil, H. Jafarkhani, "Linear Decentralized Estimation of Correlated Data for Power-Constrained Wireless Sensor Networks," *IEEE Trans. Signal Process.,* vol. 60, no. 11, pp. 6003-6016, Nov. 2012.

[54] J. Fang, H. Li, Z. Chen and Y. Gong, "Joint Precoder Design for Distributed Transmission of Correlated Sources in Sensor Networks,", *IEEE Trans. Wireless Commun.,* vol. 12, no. 6, pp. 2918-2929, June 2013.

[55] Y. Liu, J. Li, X. Lu and C. Yuen, "Design to Optimize MI for Centralized Wireless Sensor Network", available online: http://arxiv.org/abs/1412.3448.

[56] Y. Liu, J. Li, and X. Lu, "Transceiver Design for Clustered Wireless Sensor Networks — Towards SNR Maximization," available online: http://arxiv.org/abs/1504.05311.

[57] Y. Liu, J. Li, X. Lu and C. Yuen, "Optimal Linear Precoding and Postcoding for MIMO Multi-Sensor Noisy Observation Problem,", *IEEE International Conference on Communications(ICC)*, Sydney , Jun. 2014.

[58] S. Serbetli and A. Yener, “Transceiver optimization for multiuser MIMO systems,” *IEEE Trans. Signal Process.*, vol. 52, pp. 214-226, Jan. 2004.

[59] Y. Rong, X. Tang, and Y. Hua, “A unified framework for optimizing linear non-regenerative multicarrier MIMO Relay Communication Systems", *IEEE Trans. Signal Processing,* vol. 57, pp. 4837-4851, Dec. 2009.

[60] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, “An Iteratively Weighted MMSE Approcah to Distributed Sum-Utility Maximization for a MIMO Iterfering Broadcast Channel,” *IEEE Trans. Signal Process.,* vol. 59, no. 9, pp. 4331-4340, Dec. 2011.

[61] M. R. A. Khandaker and Y. Rong, “Joint transceiver Optimization for multiuser MIMO relay communication systems", *IEEE Trans. Signal Processing,* vol. 60, pp. 5977-5986, Nov. 2012.

[62] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

[63] D. Bertsekas, *Nonlinear Programming, 2nd ed*, Belmont, MA:Athena Scientific, 1999.

[64] L. Grippo, and M. Sciandrone, “On the convergence of the block nonlinear Gauss-Seidel method under convex constraints,” *Operations Research Letters,* vol. 26, pp. 127-136, 2000.

[65] P. Tseng, “Convergence of a block coordinate descent method for nondifferentiable minimization,” *Journ. of Optim. Theory and Appl.*,vol. 109, no. 2, pp. 475-494, June 2001.

[66] M.J.D. Powell, “On search direction for minimization algorithms,” *Math. Programming*, vol. 4, no. 1, pp. 193-201 1973.

[67] T. P. Minka. *Old and New Matrix Algebra Useful for Statistics*, Notes, December 2000.

[68] S. Boyd, and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

[69] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming (web page and software)," [Online]. Available: http://cvxr.com/cvx Apr. 2010

[70] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications,* vol. 24, no. 8, pp. 1439-1451, 2006.

[71] Y. Liu, J. Li, X. Lu and C. Yuen, "Joint transceiver design for wireless sensor networks through block coordinate descent optimization," available online: http://arxiv.org/abs/1409.7122.

[72] Y. Liu, J. Li, and X. Lu "Multi-terminal joint transceiver design for MIMO systems with contaminated source and individual power constraint," *IEEE Intl Symposium on Information Theory(ISIT)*, Honolulu, Jun. 2014.

[73] S. S. Christensen, R. Argawal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.,* vol. 7, no. 12, pp. 1-7, Dec. 2008.

[74] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585-595, Nov.-Dec. 1999.

[75] S.-Y. Chung, "On the construction of some capacity-approaching coding schemes," Ph.D. dissertation, Dept. EECS, Massachusetts Institute of Technology, 2000.

[76] Y. Hu, J. Garcia-Frias, and M. Lamarca, "Analog joint source-channel coding using non-linear curves and MMSE decoding," *IEEE Trans. Commun.,* vol. 59, no. 11, pp. 3016-3026, Nov. 2011.

[77] T. M. Cover and J. A. Thomas *Elements of Information Theory, 2nd ed.,* John Wiley & Sons, 2006.

[78] A. Hjorungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications.* Cambridge University Press, 2011.

[79] I. Polik and T. Terlaky, *"Interior Point Methods for Nonlinear Optimization,"* in *Nonlinear Optimization, 1st edition,* G. Di Pillo, F. Schoen, editors. Springer, 2010.

[80] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Comput. Networks*, vol. 38, pp. 393-422, Mar. 2002.

[81] K. Zhang, X. R. Li, P. Zhang, and H. Li, "Optimal linear estimation fusionâĂŤPart VI: Sensor data compression," in *Proc. Int. Conf. Information Fusion*, Queensland, Australia, 2003.

[82] J. Fang and H. Li, "Power constrained distributed estimation with correlated sensor data,"*IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 3292-3297, Aug. 2009.

[83] A. Charnes and W. W. Cooper, "Programming with linear fractional functionals," *Naval Res. Logist. Quarter.*, vol. 9, pp. 181-186, 1962.

[84] A. De Maio, Y. Huang, D. P. Palomar, S. Zhang, and A. Farina, "Fractional QCQP with applications in ML steering direction estimation for radar detection," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 172-185, Jan. 2011.

[85] C. Jeong, I. Kim, and D. I. Kim, "Joint secure beamforming design at the source and the relay for an amplify-and-forward MIMO untrusted relay system," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 310-325, Jan. 2012.

[86] W. Ai, Y. Huang and S. Zhang, "New Results on Hermitian Matrix Rank-One Decomposition," *Math. Program.*, vol.128, pp.253-283, 2009.

[87] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664-678, Feb. 2010.

[88] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol.27, no. 3, pp.20-34, 2010.

[89] A. Hjorungnes, *Complex-Valued Matrix Derivatives with Applications in Signal Processing and Communications*, Cambridge University Press, 2011.

[90] D. P. Bertsekas, *Convex Optimization Theory*, Athena Scientific, 2009.

[91] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.

[92] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Univ. Press, 1985.

[93] I. Polik and T. Terlaky, "Interior Point Methods for Nonlinear Optimization", in *Nonlinear Optimization*, 1st edition, G. Di Pillo, F. Schoen, editors. Springer, 2010.

[94] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, ser. MPSSIAM Series on Optimization, Philadelphia, PA: SIAM, 2001.

Vita

Yang Liu received the B.Eng. and M.Eng. degrees in Electrical Engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2007 and 2010 respectively.

Since 2010, he has been working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at Lehigh University under supervision of Prof. Tiffany Jing Li.

Yang Liu's research interests includes channel coding, signal estimation, beamforming and optimization theory in signal processing.

Publications List

Journal

1. **Yang Liu**, Jing Li, Xuanxuan Lu, Yuen Chau and Jun Wu, "A Family of Chaotic Pure Analog Coding Schemes Based on Baker's Map Function", *EURASIP Journal on Advances in Signal Processing,* to appear.

2. **Yang Liu**, Jing Li, Xuanxuan Lu and Yuen Chau, "Joint Transceiver Design for

Wireless Sensor Networks through Block Coordinate Descent Optimization," submitted to *IEEE Trans. Signal Process.,* Nov. 2014, revised in June 2015.

3. **Yang Liu**, Jing Li, Xuanxuan Lu and Yuen Chau, "Beamforming Design towards Mutual Information Maximization for Centralized Wireless Sensor Network," submitted to *IEEE Trans. on Vehicular Tech.* June 2015.

4. **Yang Liu**, Jing Li and Xuanxuan Lu, "Transceiver Design for Clustered Wireless Sensor Networks — Towards SNR Maximization," submitted to *IEEE Trans. Signal Process.,* Apr. 2015.

5. Jing Li, Kai Xie, **Yang Liu**, Zhiping Shi, Yong-Liang Guan, Boon Chong Ng, "Iterative Probabilistic Inference as a Nonlinear Dynamical System: State Dynamics and Z-Crease Phenomenon", *Applied Mathematics—Special Issue on Non-linear Systems*, July 2013

Conference

1. **Yang Liu**, Jing Li, Xuanxuan Lu, "Multi-Terminal Joint Transceiver Design for MIMO Systems with Contaminated Source and Individual Power Constraint", *IEEE International Symposium on Information Theory(ISIT)*, Honolulu, Jun. 2014.

2. Xuanxuan Lu, Jing Li, **Yang Liu**, "A Parametric Approach to Optimal Soft Signal Relaying in Wireless Parallel-Relay Systems", *IEEE International Conference on Acoustics, Speech, and Signal Processing(ICASSP)*, Florence, May 2014.

3. Mao Yan, **Yang Liu**, Jing Li, Qingchun Chen, "Maximizing End-to-End Capacity of MIMO Nonregenerative Relay Networks Via Time Scheduling and Subchannel

Pairing", *IEEE International Conference on Communications(ICC)*, Sydney , Jun. 2014.

4. **Yang Liu**, Jing Li, Xuanxuan Lu, Chau Yuen, "Optimal Linear Precoding and Postcoding for MIMO Multi-Sensor Noisy Observation", *IEEE International Conference on Communications(ICC)*, Sydney , Jun. 2014.

5. Xuanxuan Lu, Jing Li, **Yang Liu**, "Soft-Encoding Distributed Coding for Parallel Relay Systems", *IEEE International Symposium on Information Theory(ISIT)*, Istanbul, Jun. 2013.

6. Xuanxuan Lu, Jing Li, **Yang Liu**, "A Novel SISO Trellis Strategy for Relaying Distorted Signals in Wireless Networks", *IEEE International Conference on Communications(ICC)*, Budapest, Jun. 2013

7. **Yang Liu**, Xuanxuan Lu, Jing Li, "Reliable Signal Transmission in Wireless Sensor Networks with Zero Bandwidth Expansion", IEEE International Conference on Communications(ICC), Budapest, Jun. 2013.

8. Kai Xie, Jing Li, **Yang Liu**, "Linear Analog Codes: The Good and the Bad", Conference on Information Sciences and Systems(CISS), Princeton, Mar. 2012.

9. Xuanxuan Lu, Jing Li, **Yang Liu**, Xingkai Bao, "A New Forwarding Strategy for Wireless Relay Channels: Analog-Encode-and-Forward (AEF)", *Conference on Information Sciences and Systems(CISS)*, Princeton, Mar. 2012.

10. **Yang Liu**, Jing Li, Kai Xie, Jia Hou, "Wireless Data Sensing and Transmission through Analog Codes", *Conference on Information Sciences and Systems(CISS)*, Princeton, Mar. 2012.

11. **Yang Liu**, Jing Li, Kai Xie, “Analysis of Linear Channel Codes with Continuous Code Space”, Conference on Information Sciences and Systems(CISS), Princeton, Mar. 2012.

12. **Yang Liu**, Jing Li, Kai Xie, “Efficient Image Transmission through Analog Error Correction”, *IEEE International Workshop on Multimedia Signal Processing*, Hangzhou, Oct. 2011.

13. **Yang Liu**, Zhisong Bie, Kai Niu, Zhiqiang He,“Optimization of Constellation Labeling Using Genetic Algorithm”, *IET International Communication Conference on Wireless Mobile and Computing*, Shanghai, Dec. 2009.